

Solar energy is a key to development

Earlier this year the Commission of the European Communities held a meeting in Varese, Italy, at which scientists and technologists of both developed and developing world discussed the potential for solar energy in the latter. This meeting had been preceded by five regional seminars in Nairobi, Bamako (Mali), Amman, Caracas and New Delhi. The papers of the conference have recently been published (*Solar Energy for Development*, Martinus Nijhoff, The Hague) and raise some interesting questions concerning a science and technology whose relevance to development is very close indeed.

In the developed world, it is going to be a very long haul for solar energy to supply a significant fraction of our needs — estimates generally range around a contribution of a few per cent at most by the year 2000. This slow rate of progress is less a reflection on lack of scientific and technological skills than on the enormous built-in momentum that modern industrialised society has which strongly resists major changes of course. The picture is very different in the developing world. There solar energy (in the form of direct heating, biomass and wind) is, and always has been, the dominant energy input, particularly in rural environments. In Africa, for instance, even by 1985 not one in ten of all villages will have an electric power supply.

In many ways this dependence on the sun should be regarded as a strength for the future, which should not lightly be yielded up to the blandishments of other sources, for the poorest people of the world can ill afford to get caught in the spiral of energy price increases. And in the long run, as Sir George Porter pointed out to the meeting, if and when the technology of chemical fuels is developed, it is the barren deserts of the world which could well become the ideal place for growing these fuels — and hitherto resource-poor countries could become major energy exporters. There is also a fairly hardnosed political reason why developed countries would be encouraging developing countries to think solar — that it might just head off some nations from thinking

nuclear with all the complex ramifications that would bring.

All of this clearly adds up to a compelling case for the vigorous pursuit of more research and development into uses of solar energy. Favoured as many developing countries may be in their use of cheap uninterrupted energy from the sun, they cannot develop much further without demanding more energy from somewhere.

If solar energy is to be even more widely used in the developing world, two major obstacles will have to be overcome. The first is the lack of status for the subject. Someone who designs a new solar water heater or crop dryer or desalination plant may contribute far more to development than someone who works on an exotic theoretical problem in an academic environment. And yet the latter will get more acclaim, particularly from an international community. It is no surprise that the developing world only started to show strong interest in advancing solar science and technology when developed countries (and particularly the United States) went in for it in a big way. But even now, it is not at all clear that adequate intellectual resources are being devoted to the problems.

The second obstacle is the understandable feeling in most developing countries that research and development, if done at all, has to be done domestically. Collaborative regional centres of research, where they exist at all, often get the problems that have lowest priority. But the sort of questions raised in solar energy research rarely fit snugly to national boundaries, and there is every reason to pool resources over quite wide regions. Slowly this is being realised in Africa and Latin America, but there is much scope still for more effective collaboration, if only at the level of the exchange of information.

Solar energy could be a large ingredient in the economic advancement of the developing world. It offers a marvellous opportunity for the use of indigenous skills. Will the developing world take up the challenge? □

Environmental regulations under attack

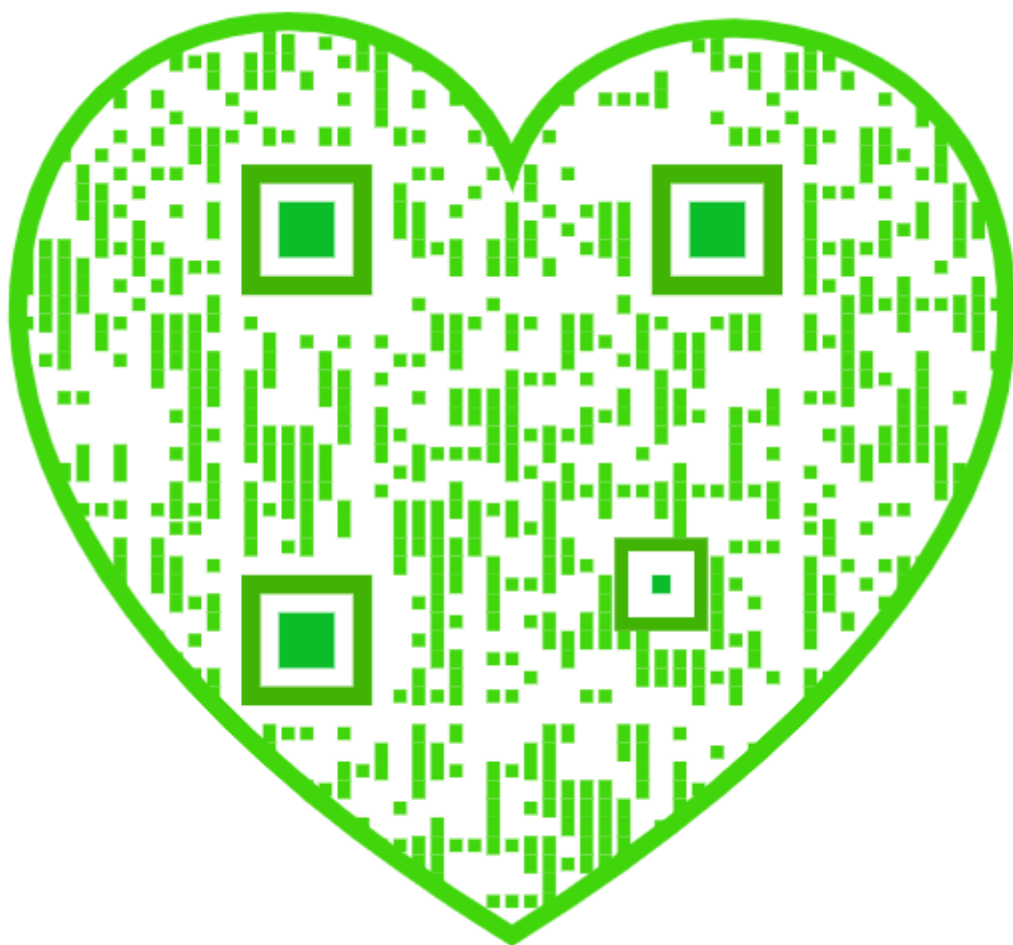
The end of a decade is inevitably a period of taking stock, looking back at the impact of past actions and forward to their implications. Ten years ago the environmental attack on the unwelcome side-effects of rapid industrialisation was at its height. Environmental protests became an increasingly common event across the western world. Demonstrators' concerns ranged from the effects of pesticides on wildlife to the carcinogenic hazards of the workplace. In the US, Congress responded with legislation such as the Natural Environmental Protection Act of 1970, and the Occupational Safety and Health Act of the same year. And the air rang with 'end of an era' quotations about how society was learning from the error of its past ways.

We may be a few months early in bringing down the curtain on the 1970s. But in recent weeks, political events in the US have begun to overtake the calendar in signalling the end to the spirit which has pervaded much of the past decade — already referred to by some as the 'golden age of environmentalism'. Bold commitments to guaranteeing a clean and safe environment, with the federal government being given a prime responsibility for achieving each, are giving way to increasing calls to pass this

responsibility back to the marketplace and to the private sector. The effects of regulation on energy supply and on technological innovation are usually mentioned as the most important reasons for this shift. But both are backed by a growing anti-government tendency in US politics, with Washington's role as chief regulator placed squarely in the stocks.

Over this next three weeks we shall be running a series of news features on this trend and its implications for the future. In this issue we examine (page 168) the growing conflict between environmental and energy demands. Next week we shall look at the industrial response to recent moves in the fields of occupational health and safety. Finally we shall examine how, in the international nuclear energy domain, US attempts to control the spread of nuclear technology have met with, at best, only limited success.

We do not wish to speculate too much on what the next decade has in store. But the movements of the past decade will occupy an important chapter in the history of the relationship between science and politics. The movements of the next are unlikely to be any less significant. □



The end of the age of Aquarius?

Concerns about energy and inflation, linked to a conservative trend in US politics, are beginning to eclipse the moral commitment to a clean and safe environment characteristic of the past decade. In the first of three articles on how this has happened and what may follow, **David Dickson** describes how the energy issue has become a major threat to US environmental movements.



"Man can survive pollution; environmentalists can relax about that. The question is whether he can survive regulation." These words headed a full-page advertisement in the US national press last week for a new book, *Restoring the American Dream*, by best-selling author Robert J Ringer.

The sentiments are not new. Indeed such arguments have been frequently heard since concerns about the effects of pesticides on bird-life ushered in the environmental movement in the mid-1960s. And they have provided a constant back-drop to a decade of legislative measures designed to protect the US environment against the worst excesses of technological growth.

In the past few months, however, concerns about growing inflation and the security of future energy supplies have both helped to generate an anti-regulation bandwagon that is demanding a major re-think of this legislation and its effects — as well as laying the ground-work for a substantially different approach to environmental regulation in the 1980s.

As far as the administration is concerned, this will mean less emphasis on the government as a regulator of industrial activity, and greater emphasis on the market-place as the matrix for 'rational' decision-making on regulatory activities.

For the environmental movement, it means a shift from the essentially defensive posture which the movement has been adopting in recent years, capitalising on the gains made earlier in the decade, to one in which urgent action is now needed in most fields (except perhaps that of nuclear energy) to prevent the threat of a complete rout.

Two main factors have fuelled this anti-regulatory tendency. The first has been the argument that inflationary problems are directly linked to a decline in national productivity (which fell by an annual equivalent rate of 5.9 per cent in the second quarter of 1979); and that this is largely linked to a decline in technological

innovation, itself due partly to increased environmental and health regulation.

"What started as a regulatory snowball ten years ago has become an avalanche, and innovation is one of its victims," Dr H D Doan, ex-president of the Dow Chemical Company, said at a meeting of the American Chemical Society in Washington last week. "George Orwell may not have been accurate in his predictions; but the regulators tend to confirm his foresight."

The second factor which has played a key role is the pressure on US politicians, following the Iranian oil crisis earlier this year, to be seen to be doing something about guaranteeing future energy supplies — if necessary at the expense of environmental considerations.

Two developments in particular have led to a qualitative rise in the energy *versus* environment debate. The first has been the push for a massive programme to stimulate the production of synthetic fuels as a substitute for imported oil, and the renewed emphasis on the use of coal, both of which, if not properly handled, could have a significant impact on factors such as the build-up of atmospheric CO₂.

The second factor, which many environmentalists fear could be even more damaging, is the president's decision to establish a top-level Energy Mobilisation Board. Of particular concern are the strong powers which many congressmen would like this board to have to over-ride federal and state environmental legislation in order to speed up major energy programmes.

The idea that there are necessary trade-offs between energy demands and environmental considerations is not new. Until relatively recently, however, the administration has sought to contain such debates within a framework of cost-benefit analysis, using economic pay-off as the means of weighing up the merits of one course of action against another.

Such analysis, for example, has been behind some recent decisions to relax

certain aspects of clean air legislation. For example, many power companies are being squeezed between new limits on sulphur dioxide emissions, and the president's call to reduce oil consumption by half by 1990, largely through the use of coal; and cost-benefit analysis is being used as a prime tool to tackle such conflicts.

In the past few months, however, this approach has been overtaken by a more overtly political drive aimed at circumscribing the government's role in a range of activities. One of the first manifestations of this trend was California's decision last year to halve its property taxes, the famous Proposition 13; more recently, government regulators have become a prime target for political abuse.

According to a survey published in March by Opinion Research Corporation, for example, more than half of the general public feel that the costs of government regulation exceed the benefits. And by a margin of two to one, those interviewed said that there was a need for less rather than more government regulation.

With elections coming up next year, congressmen from both parties are eagerly responding to what they see as popular demands for less regulation — in many cases outstripping the administration's economists in their anti-regulatory fervour. In the last two weeks, for example, this tendency has become clear in four separate congressional actions.

The first was an amendment by Senator Dale Bumpers (Dem. Arkansas) to a bill on the federal court system, requiring a federal agency to demonstrate the validity of a particular set of regulations in court every time that they are challenged. The amendment thus challenging the "presumption of validity" that now covers agency rules when challenged in court proceedings, making it easier for courts to overturn federal regulatory decisions.

The second decision was taken during a debate last Tuesday on a strip-mining bill, which would allow states to disregard regulations laid down by the Department of the Interior about making good disturbed landscape. On the same day, the Senate also voted to allow the completion of the Tellico Dam, thus in effect agreeing to the extinction of the snail darter, whose status as an "endangered species" had previously led to a Supreme Court ruling that the dam should not be finished.

Perhaps the most significant decision, however, was taken by the House Commerce Committee. In discussing the role of the President's proposed Energy Mobilisation Board, the committee agreed that the board should have considerably stronger powers than the president had proposed, enabling it to waive environmental regulations.

Despite intense efforts by environmental



lobbying groups, the committee approved a bill which would allow the board to recommend that the president waive any federal, state or local law blocking a high-priority energy project.

Environmentalists are supporting an alternative version of the bill under which, as originally proposed by the administration, the board would be merely able to speed-up procedural aspects of project approval; and one sympathetic congressman, Representative Tim Wirth (Dem. Colorado), has promised a "huge, huge battle" when the Commerce committee bill is debated by the full House.

Critics of the stronger version of the bill have been particularly concerned about the apparent support that this version is now receiving from the White House (even though administration officials claim that their support is at present essentially a bargaining device).

"If we are not careful, the Energy Mobilisation Board could threaten fifteen years of environmental work," says David Maselli, energy policy director for Friends of the Earth. He points to the low priority given to environmental considerations in the energy proposals with which President Carter emerged from his recent seclusion at Camp David; "it was close to the bottom line — so close to a 'worst case' guess about the outcome of Carter's deliberations that many environmentalists were genuinely surprised," he says.

Maselli points out that environmental groups have not lost out on everything in recent legislative debates. They have had considerable success, for example, in slowing down the projected growth of the administration's proposed programme for synthetic fuels — an issue against which "we are standing as hard as we can" says Herb Epstein of the Solar Lobby — and have won acceptance for numerous solar energy and conservation measures.

But no-one denies that there are some major battles ahead. "I think it is going to be a hellish struggle, even to preserve the

gains that we have won," says Epstein. And others are already saying that the 1970s may be looked back on as "the golden age of environmentalism."

Speaking at the ACS meeting last week, Dr Frank Press, director of the Office of Science and Technology Policy, indicated a shift in administration thinking when he said that he saw a "new philosophy" being pursued. This was one that recognised that the desired goals of environmental, health and safety regulation might best be achieved through incentive-oriented, rather than a direct federal regulation, approach.

"Such an approach, rather than commanding that specific technologies be applied to achieve the specific standards, allows the industry to choose its own paths to achieving the same end results," he said. And he gave as an example the "bubble approach" which the Environmental Protection Agency had introduced to air pollution, in which a plant can set its own effluent release from each point source as long as the total effluent release complies with a federal standard set for the entire complex.

"One of the important things to recognise regarding this approach to regulation is that it provides new incentives for industry to innovate. It takes some of the repressiveness and uncertainty out of regulation, and relies more on market forces to achieve the desired goals in environmental quality, health and safety."

The need to allow an increased, if not dominant, role for market forces is also emphasised in a report *Energy: The Next Twenty Years*, sponsored by the Ford Foundation and published in Washington last week by Resources for the Future.

"Air pollution control policy should concentrate on providing incentives for making progress toward cleaner air in a way that is cost-effective over time," says the report, produced by a group of academics and industrialists. Attempts to define rigid "safe" exposure levels are both

scientifically and administratively unsound, the report recommends moving towards "a different philosophy" of pollution control, deemphasising rigid deadlines and standards, in which market processes determine the details of ways to meet particular environmental goals.

The new shift is also likely to have an effect on the academic environmental science community. "The new climate for environmental regulation makes the need for accurate scientific data even more important than before; in particular there will be a need for better epidemiological data, looking at the real effects both of pollutants and of environmental regulations," says Dr John E Cantlon of Michigan State University, chairman of the National Academy of Sciences' Environmental Studies Board.

"I do not think that all the census work that I have seen indicates that people are willing to give up clean air or clean water. But they may be more willing than before to ask the question of whether the cost-effectiveness level of certain regulations is defensible," says Dr Cantlon.

Finally the environmental movement is beginning to realise that the strategies which it used so successfully in the early 1970s to achieve significant legislative victories may no longer be appropriate for the form of the debate in the 1980s. And that it may be necessary to form new coalitions — particularly with labour or other political groups — if it wants to remain effective.

"One of the weaknesses of the environmental movement is that it has never been able to deal properly with economic issues. But recent moves by the administration on issues such as synthetic fuels mean that the stakes have been raised, and we must deepen both our understanding and our political linkages if we wish to remain effective," says Gail Daneker of Environmentalists for Full Employment.

What effect environmental issues have on next year's presidential election remains to be seen. Certainly many of those who actively supported President Carter earlier in his administration are growing increasingly sceptical. "The one thing worse than having a public figure against you is having a powerful public figure who used to be on your side against you," says David Maselli, referring in particular to the administration's apparent support for a strong Energy Mobilisation Board.

But few see the issue merely in terms of personalities — or of presidential politics. Rather it is a response to an economic and political climate fundamentally different from that which saw the dawning of an "Age of Aquarius" in the late 1960s. With demands for a decreased government role in regulation, and an increased emphasis on market-place economics, moral commitments no longer carry much political capital. □

Next week: health and innovation

ICSU to raise human rights issue at 50 meetings

This summer, a new committee on the 'Safeguard of the Pursuit of Science' of the International Council of Scientific Unions delivered its first report to the ICSU general assembly. The purpose of the committee is to investigate and record cases of scientists unwarrantedly dismissed from their posts or otherwise hampered from pursuing their careers. During its first year's work, the Committee had been apprised of around a dozen such cases, in countries ranging across the whole political spectrum.

This low count, however, does not represent a sudden worldwide liberalization of hitherto oppressive regimes. Rather does it reflect the extremely unobtrusive manner in which the Committee has commenced its activities. For example, not a single case of *berufsverbot* was notified to the Committee. This does not, alas, mean that after all, there are no West German scientists suffering for their allegedly left-wing views. A few days ago, Dr Andreas Dress of Bielefeld University, a leading campaigner against *berufsverbot* in science, told *Nature* that he had not heard of the Committee.

At the 1978 General Assembly of ICSU, when the Committee's founding statement was finally approved (this is a most diplomatically worded document, incidentally, avoiding all such emotive terms as 'human rights'), it was decided that the work of the Committee was an internal affair of ICSU. Accordingly, the statement was never officially issued to the general or scientific press, and only reached *Nature* by informal means almost a year later. Information about the Committee, it was decided, should be disseminated via the usual channels — the national representatives to ICSU (normally national academies of science or their equivalents).

This inevitably suggests a Catch-22 situation. An academy which was actively or passively endorsing the kind of discrimination and repression which it is the Committee's brief to monitor, would hardly be expected to pass on information to its potential victims as to how to obtain a measure of redress. The Committee, however, is less pessimistic. Its Secretary, Professor John Humphrey of the Royal Post-Graduate Medical School, London, told *Nature* that "We start with the premise that the national representatives are no happier than we are that their scientists are being removed from academic life". He admitted, however, that one cannot assume that in such countries the scientific establishment would disseminate news of the Committee very vigorously.

To spread the word more rapidly, therefore, the Committee has decided on a limited publicity campaign. They will display and distribute the founding statement and the forms for reporting

abuses at some 50 international scientific meetings in different countries and different disciplines. Visiting scientists, who have not yet heard of the committee through their own national organizations will, it is hoped, be able to publicise the committee when they return home.

The committee's main brief is to investigate and record abuses of a scientist's right to work in his profession. ICSU is a prestigious body, and it is hoped that the moral censure it can impose will have a considerable restraining influence. In the case of highly oppressive or arbitrary regimes, where citizens are likely to disappear without trace, it has been suggested that the committee should play a more active role. Instead of waiting for cases to be reported, the committee could keep a list of all known scientists working in the country, and take as a warning signal any untoward silences or absences from

conferences at which they might be expected to appear.

Moreover, in the coming year, the Committee plans to extend its activity to a more general impediment to the free pursuit of science — the vexed question of official secrecy. A working part is to meet to discuss "the losses to science that arise from all kinds of restriction on communication". A total disappearance of the concept of 'classified information' is not envisaged, Professor Humphrey stressed. Rather there would be discussions of how to remove unnecessary restrictions, whether scientists possess any special status, the significance of boycott to the growth of science and so on. It would be, he added "a Pugwash kind of thing, dealing with the impact of restriction of communication of world science, rather than with the rights of scientists."

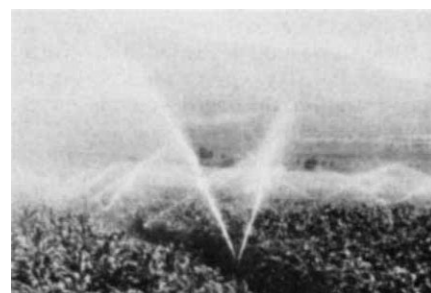
Vera Rich

Bulgaria — proud of science on the farm

"If we say that land is unproductive, it is we who are to blame. It simply means that we have not yet discovered the most rational type of crops for it". So Professor Milko Iordanov, a leading Bulgarian agronomist, told journalists recently. Appropriately, he was speaking at one of the show-pieces of Bulgaria's latest development in agricultural planning and research, the "Georgi Dimitrov agro-industrial complex", in the Rhodope foot-hills.

Agro-industrial complexes, explained Professor Iordanov, are essentially a further stage in the process of amalgamation which, "from 9 September, 1945" onwards, led to the formation of large collectives and cooperatives out of peasant small-holdings. By 1970, he said, it was clear that even cooperatives of 6000 to 12,000 ha still did not permit the "optimal use" of modern technology. Accordingly, the Party adopted as a special aim, the creation of agro-industrial complexes, in which several villages are combined into a single production unit, with its own food-processing plants and agricultural research institutes. Within each complex, said Professor Iordanov, each farm still retains considerable autonomy. The complex, however, is the "general leader", in charge of capital investment and overall policy. Typically, each complex will produce a wide range of crops. At the "George Dimitrov", these change with altitude from orchard fruit, through grapes, tomatoes, hazel nuts and almonds, to potatoes, and finally, sheep grazing.

The research institutes established, in Professor Iordanov's words "for complexes working on a large scale" are subsidised by national funds. Their results and experience are passed on free of charge to any other complexes which are interested although, should experts as well



Irrigation in the Rhodope

as information be required, it appears that some kind of bilateral contract is established between the complexes, National funding, Professor Iordanov intimated, is given "according to progress", and the research plans are in a considerable extent tied in to the national economic plans. Thus, at the "Georgi Dimitrov", work is going forward not only on specific agricultural problems but also the more general themes of saving energy and material resources. Already, said Professor Iordanov, they have succeeded, by combining a modified American "tiler" cultivator and a Soviet tractor, to obtain the same productivity with a considerable saving of metal.

For the most part, however, research is concentrated on crop investment and the selection of the most productive strains. There is considerable emphasis on "technologies for abroad". Libya already has orchards grown from Bulgarian seeds, and is sending back fruit under a compensation deal. Within Comecon, which has a common research programme on crop selection, the Bulgarian agro-industrial complexes likewise play a major role — indeed, Professor Iordanov admitted — he himself happens to be coordinator for the whole programme.

Vera Rich

news in brief

CIA told to reveal university research links: The US Central Intelligence Agency is fighting a court order to reveal the names of universities at which it supported research into the effects of drugs on human behaviour during the late 1960s, under a programme code-named M K-ULTRA. A Washington judge has given the agency until the beginning of next month to disclose the names of the universities involved. Judge Louis Oberdorfer gave this ruling following a Freedom of Information Act suit which had been filed by John Cary Sims and Dr Sidney M Wolfe of the Health Research Group. However at the same time, the judge also suggested that the CIA might be able to argue that releasing the names of the universities would involve a "clearly unwarranted invasion of privacy" of the individual research workers. He also said that the names of the universities themselves might be reclassified as information which it was "in the interest of national defense or foreign policy" not to reveal. John Sims of HRG said last week that the names of a number of the universities involved in the experiments were already known, but that it is important to establish the basic principle. "Information about the MK-ULTRA programme raises the question of what right students and staff have to know about the research being conducted on their campus. Furthermore institutions which have already been identified as working with the CIA have been forced to face publicly the issue of what their relationship with the agency should be."

A spokesman for the agency said last week that it would attempt to persuade the judge that it was still entitled to withhold the names of the universities involved in the programme. It is understood that the director of the CIA, Admiral Stansfield Turner, has already written to all the universities concerned asking if they object to having their names revealed; and that some have replied that they would prefer their involvement with the agency not to become public knowledge.

Thibault named head of reorganised CNRS: Professor Charles Thibault, 60, has been appointed to the newly created post of president of France's reorganised Centre National de la Recherche Scientifique. Thibault, a specialist in reproductive physiology is director of research of INRA, the French agronomy research institute, and has held the chair in the physiology of reproduction at the University of Paris since 1967. He is known for his work on *in vitro* mammalian fertilisation and is a firm advocate of research oriented towards applications. His appointment is expected to accelerate the government's controversial drive to place French research on a more commercially profitable basis.

6000 demonstrate against Scottish nuclear power:

SCRAM, the Scottish Campaign to Resist the Atomic Menace, a coalition of 40 anti-nuclear groups, organised one of Edinburgh's largest demonstrations last Saturday as part of a campaign against the development of nuclear power in Scotland. Speakers outlined the dangers of nuclear power generation at each step of the process from mining to waste disposal. Uranium mining, scheduled for Orkney, was singled out for special emphasis because of the large scale environmental disruption that is known to be caused by a process requiring tons of ore to be mined to produce ounces of uranium. In addition the hazards of radon releases associated with mining and toxic chemical wastes associated with the extraction process were emphasised.

Other speakers drew attention to recent controversy surrounding the hazards of low level radiation expected to be

associated with the operation of the projected plant at Torness and the drilling for prospective waste disposal sites in northwest Scotland and Ireland was strongly opposed. Local feeling against nuclear dumping runs high and one speaker drew loud applause with the comment: "If it's so safe, why don't they do their drilling in London?" The demonstration attracted trade union representation for the first time in the UK. Mr Mick McGahey, president of the Scottish miners told the rally that nuclear power should not be developed until a comprehensive energy policy is evolved and until strict safety standards could be guaranteed. Coverage of the demonstration was enlivened by the presence of actress Julie Christie (below left) who marched with the Parents Against Torness group. Christie is opposed to the 15 waste disposal drilling sites, one of which is scheduled for Powys, Wales, where she lives.

ESA moves to commercialise Ariane space launcher:

The European Space Agency is putting the finishing touches to a French plan to create an industry-controlled corporation to administer Ariane, the European space launcher. Tentatively called 'Ariane-Espace' the new organisation will have control over manufacture and launching of all future spacecraft with government represented only by a minority control. Raymond Orye, head of the Ariane programme explains the move in terms of the usual industry-government relationship. "Once a product is developed it should be handed over to industry for use" says Orye. On the question of industrial control of the corporation Orye explains that the capital risks involved means that "industry needs liberty in order to get maximum return on their investment."

The commercial applications are largely in telecommunications, with 75% of future missions having telecommunications applications. But other exploitable areas are expected to be in Earth resources surveying and in meteorology. At present ESA is in advance of the US in turning over the results of space research to private development. The 11 member states are in broad agreement on the plan and the details on contractual and legal matters, price policy and launch control are being debated in a series of closed Council meetings. The next meeting is scheduled for 10-12 October in Paris.

Traditional Chinese medicine takes new direction: Following a recent document issued by the Party Central Committee on applying the skill and experience of experts, the Beijing (Peking) Health Department has organised its research work in traditional medicines on a new basis. Thirty nine experts have been chosen to take on apprentices or assistants for clinical practice to learn particular skills in treating certain diseases.

In the meantime new books on the clinical practices of individual experts in traditional medicines have been compiled by a team of 40 experts working over the past four years. These include *Liu Fengwu's Experience in Gynaecology*, *Selection of the Medical Experience of Guan Yubo* (an expert in hepatitis) and *A Collection of Clinical Experience of Zhao Bingnan* (a noted dermatologist) with books on surgery and paediatrics to be issued soon. Work of this sort was started in 1959 but was interrupted by the Cultural Revolution when many academic records were lost and some of the experts were persecuted badly. In 1975, the work was renewed at the bidding of Premier Zhou Enlai.



Asbestos: killer dust, says book

A new book on asbestos claims half a million will die in the UK of asbestos-induced cancer. **Alastair Hay reports**

Five hundred thousand people will die from asbestos-related diseases in the UK in the next forty years according to a book published recently by the British Society for Social Responsibility in Science. In 'Asbestos Killer Dust' the author, Dr Alan Dalton, says that a 1978 US Government Report on asbestos-related cancer deaths makes possible the prediction that a "real epidemic of asbestos related diseases" is about to occur in Britain.

The US report, published by the National Institute of Environmental Health Sciences (NIEHS) estimated that over 2 million American workers will die of cancer as a result of occupational exposure to asbestos. Thus, says Dalton, in Britain — with a population roughly a quarter of that in the US — it can be assumed that asbestos will kill some half million people.

Is the future that grim? It could be, says Dalton, if the recent reports from the US are correct and if Britain does nothing about the current standards regulating asbestos fibre concentrations in the workplace. The standards — based on the number of fibres per millilitre of inhaled air — differ for the various types of asbestos used. Current UK recommendations are 0.2 fibres ml^{-1} for blue (crocidolite) asbestos and 2.0 fibres ml^{-1} for white (chrysotile), brown (amosite) and any other fibre.

According to Dalton the present British standards allow much too high a concentration, and far from being safe present a positive danger to health. Citing a 1976 US National Institute for Occupational Safety and Health (NIOSH) report that there is no safe level for asbestos, the author questions the justification for the British standard.

Dalton points out that, in its report, NIOSH recommended a new US standard of 0.1 fibre ml^{-1} — but that even this was chosen because of practical problems in taking accurate measurements of asbestos fibre concentrations. Why then, the author asks, is the British standard for white asbestos 20 times that recommended for the US? And why he questions, does Britain persist with its two-tier standard for blue and white asbestos when NIOSH recommends one standard for both? Mesothelioma, the asbestos-related cancer, Dalton points out, is now quite clearly attributable to exposure to both types of asbestos and not just to the blue form as was originally believed.

The British standards are currently being reviewed by the Advisory Committee on Asbestos of the Health and Safety Executive (HSE), and a report stating the Executive's view on the adequacy of the standard is expected in the next 2-3 months. It has been reported that the committee will recommend that the standard for white asbestos should be

halved to 1 fibre ml^{-1} , matching Sweden's limits; that the limit for brown asbestos should be reduced to 0.5 fibre ml^{-1} ; and that blue asbestos should be banned, with the standard for its removal remaining at the present 0.2 fibres ml^{-1} .

It is quite clear, however, that the original study at the Turner and Newall factory in Rochdale in 1966 which laid the basis for the British standard was flawed. The number of asbestos-related health problems was considerably underestimated. Only workers who had completed 10 years service and were still in the company's employ were included in the study; those who had left the factory were not followed up. In addition the clinical test used to assess the effect of accumulated asbestos exposure has been criticised as being too stringent.

The test measures the prevalence of crepitations (abnormal sounds in the lung during exhalation and inhalation) and it has been suggested that this test might only identify more serious asbestos injury and miss cases with minor damage. Furthermore, there are indications that the dose/response data from the Turner study may in fact be wrong, as the dust level was actually higher than originally reported.

Although there is no official comment from UK Regulatory Authorities on the 1978 NIEHS report for asbestos related cancer deaths, criticisms have been made about it in private. For example, it is known that the HSE feels that figures for deaths in the US cannot be applied to the UK. Scientists at the HSE are known to have doubts about the validity of the assumption in the NIEHS report that the "relative risk" of developing cancer is constant throughout life. Furthermore assuming the most conservative estimate of risk in the NIEHS report it can be calculated that at least 5000 deaths from mesothelioma should be occurring annually. In 1976, however, the US figure was 1000 deaths.

Even though the HSE feel that Dalton has considerably over-estimated the dangers, the Executive does admit that mesothelioma deaths present a real problem. Figures collected by HSE's Employment Medical Advisory Service show that deaths from mesothelioma have risen inexorably since 1968. EMAS figures for 1968, 1972, 1974, 1977 are 154, 208, 230 and 323 deaths respectively. An HSE spokesman pointed out that the mesotheliomas may not have been due to asbestos; but he did confess that it was almost certain that they were. □

'Asbestos Killer Dust' by Alan J.P. Dalton. Published by the British Society for Social Responsibility in Science, 9 Poland Street, London W1V 3DG. Price £2.25

Desai's difficulties with AT

Alternative technology is no simple panacea for India writes **Anil Agarwal**

WITH the emergence of the Janata government in 1977, considerable hopes were raised both within the country and outside, that India would become a major Third World country actively using and supporting appropriate technology (AT). The need for AT was mentioned in the election manifesto of the Janata Party and, in his very first broadcast to the nation, Prime Minister Morarji Desai, had emphasised that "the big city, the big machine and big science have their place. But they cannot claim a prescriptive right to preference and dominance. We have to make the village the centre of our economic progress."

The industrial policy announced by the Janata government in December 1977 also recognised the importance of reducing unemployment and rural-urban disparities in India's future industrialisation programmes. Although India is today the tenth largest industrial producer in the world, the total number of its unemployed is almost equal to the total number of those employed by all its organised mills, factories and services. Mr George Fernandes, the Janata Industries Minister, therefore, declared that AT "will henceforth be an integral part of policy and government will ensure that this important area gets adequate attention." The Janata government, he stated, "will not favour large scale industry merely for demonstration of sophisticated skills or as monuments of foreign technology. The main thrust of the new industrial policy will be on effective promotion of cottage and small industries widely dispersed in rural areas and small towns."

AT, thus, appeared to be in early 1978 not only a sensible option for India, but also one of enjoying great support. However, over the months the initial resolve of the Janata government in favour of small-scale labour-intensive rural-oriented AT was weakened. This was, to a great extent, because of the myriad inter-party factional tensions, which virtually sapped the very will of the government to govern, but partly also because of the strong opposition from scientific and industrial establishments.

The Janata government's advocacy of AT was unfortunately accompanied by a barrage of criticism of the country's past scientific and technological activity. The critics argued that the country had not adequately benefited from its large investments in R&D institutions and that



Spinning polyester: more jobs, but does it increase dependence?

the direction of past R&D activity was too urban and big industry-oriented. Eighty per cent of Indians, they argued, live in the villages yet very few technologies have been developed to meet their needs.

This criticism obviously drew counter-criticism. Professor M.G.K. Menon, the former Chairman of the Electronics Commission and the current Director-General of the Council of Scientific and Industrial Research (CSIR), for instance, strongly argued that although research can be done to find new ways of using agricultural and animals wastes in rural areas, and so on, it is nonetheless wrong to criticise the CSIR and its institutions for neglecting this kind of research because it was not in their initial charter to do so. Secondly, some scientists even argued that the government's advocacy of rural development, appropriate technology — and, in fact, even its efforts to reorganise bodies like the CSIR (See *Nature*, vol) — betrayed an "anti-science and anti-technology belief". Professor A.K.N. Reddy of the Indian Institute of Science at Bangalore, a well-known advocate of AT, alleged in an article in an Indian economic journal that there is a "widespread suspicion that the present government, in its over-riding concern for rural problems, believes that these problems can be solved *without* science and technology." Dr Reddy stated his own belief that the generation of AT is "rarely a trivial affair — it invariably requires subtle analysis, innovative experimentation and design, and sophisticated scientific and engineering thinking." According to Reddy, developing countries interested in AT must develop their own indigenous sophisticated scientific skills and that they must conduct basic as well as applied research. To take one instance, basic studies of methanogenic bacteria can greatly contribute to the development of biogas systems.

Some senior Indian scientists and science policy experts also denounced AT as a concept that is largely being promoted by those western lobbies who want to belittle India's past efforts towards technological self-reliance and keep Third World countries permanently dependent and

backward. In what was an obvious reference to developments in India, Mr Ashok Parthasarathi, the current Secretary to the Electronics Commission and former special assistant on science and technology to Mrs Gandhi, told a seminar on the UN Conference on Science and Technology for Development (UNCSTD) held in Sweden in late 1978, that there is a strong body of opinion amongst the "nationalist segment in most Third World countries" that regards AT "as a diversionary tactic on the part of the North — diversionary in the sense that it seeks to redirect a significant part of the scientific and technological (S&T) capacity of Third World countries towards the development of these new 'rural-oriented', 'appropriate' technologies, leaving the main muscle of industrial power — the petrochemical complexes, the steel plants, the railways, the power plants — in the hands of transnational companies."

Indian scientists who share this argument point out that even if a Third World country were to make rural modernisation and agricultural development the prime goals of its development strategy, it would still need several sophisticated industrial products like cement, chemical fertilisers, pesticides, iron and steel, trucks, railway systems and electricity to make its strategy work. The technology required to produce all these products is today dominated by western multinational corporations.

The critics further challenge the value of diverting large scientific efforts towards the development of rural-oriented AT by pointing out the strong evidence that has emerged in recent years that what is known as AT, in a highly stratified society like that of India, tends to benefit the rural rich and not the rural poor. A prime AT like the development of biogas plants has been found to have benefitted only those landowners who possess the money and cattle to buy and operate such plants. Rural modernisation and eradication of poverty, the critics argue, require a political will for structural transformation within society. So, it is claimed, if a Third World country like India were to pay too much attention to AT it would not only fail to improve the lot

of its rural poor but also, in the process, become more dependent for its modern technological needs on foreign multinational corporations.

Indian supporters of AT like Mr M.M. Hoda of the Appropriate Technology Development Association in Lucknow protest that the proponents of modern western technologies, while raising fears of a strengthening of the rural elite, are conveniently forgetting that their own technological strategy will lead to the strengthening of an already powerful urban-industrial elite. And that this elite will equally oppose any structural transformation that may be required in the long run to eradicate poverty. The argument that the Janata leaders are "anti-science" appears to reflect the elitist character of the scientific community itself. It is indeed not rare to hear an Indian scientist calling the Janata leaders a pack of 'desi' (native) people, largely because they do not possess the finesse that earlier leaders like Mr Nehru and Mrs Gandhi acquired from their western educations.

AT-protagonist Reddy, however, does criticise western organisations and aid-agencies for their excessive advocacy of AT. By giving the impression that AT is something mainly suitable for the Third World and by doing precious little to develop 'appropriate' technologies for use in developed countries themselves, Reddy argues that western organisations are not just being extremely patronising but are also producing very counter-productive results with their advocacy. Their enthusiasm, he claims, has created such a strong backlash to the concept of AT in India that even Indians are now finding it difficult to advocate it.

"What India needs is both large and small-scale technologies," explained an anti-AT senior policy-maker to me, a few months before Mr Desai's government fell. "We should not be asking organisations like CSIR to redirect their efforts towards rural-oriented AT but instead be asking them to step up their research efforts to develop both small-scale rural-oriented technologies as well as those technologies that are required by modern large-scale industry." "I think," he proudly claimed, "that we have now succeeded in educating the new Janata leaders, including the Prime Minister."

That a Third World country like India needs a mix of technologies was pointed out even in the Indian national paper to UNCSTD. The paper explains that while the broad criteria for the appropriateness of technologies must be defined by a nation's development objectives (like meeting basic needs, employment-generation, self-reliance and environmental protection), the implied productivity *versus* employment problems can be resolved only by adopting a balanced mix of capital-intensive and labour-intensive technologies, of centralised and decentralised production,

and of large-scale and small-scale units.

But what exactly is it that the Janata government has tried to do in the home of AT, apart of course from speeches by Prime Minister Desai and other Janata leaders that Indian scientific research needs a different direction? The clear answer is: 'not very much'. The main attempt to go small and introduce the concept of appropriateness in technology choice was restricted to the — albeit important — industrial sector, and here too the Janata government merely tried to put more teeth into a policy that all former Indian governments have advocated since the mid-1950s to protect and promote small-scale industry. The Ministry of Industries increased the number of items exclusively reserved for production in the small-scale sector from 180 — under Mrs Gandhi's government — to over 500. And to help promote the development of small industries in neglected backward areas, it launched a scheme to set up an industry centre in each district of the country. The centres will provide small entrepreneurs with all the services and support they require under one roof, including economic investigations of a district's raw material and other resources, supply of machinery and equipment, provision of raw materials, arrangement of credit facilities, quality control and research. The scheme, which started in May 1978, has had teething difficulties, but 223 centres had been sanctioned by end of 1978.

In certain industries like cement, matchsticks, soap and textile manufacture, which can be operated on a small scale, Mr Fernandes — the Minister for Industry — strongly supported efforts to make small-scale ventures succeed. For instance, his ministry provided a 50% rebate in excise duty for cement produced from mini-cement plants (with a capacity less than 200 tonnes per day). Dr S. Chopra, Director-General of the Cement Research Institute, which has developed 20 tonne per day capacity plants, estimates that mini-cement plants should be producing about 10% of India's cement within the next five years. Mr Fernandes also strongly advocated the manufacture of cement from rice husk ash, though some scientists feel this was premature, as there are very few such plants in existence in India and the quality of their product still needs to be standardised.

Under Mr Fernandes, the Ministry of Industry also asked those large industries, many of which are subsidiaries of western multinationals, which produce such mass consumer goods as soap and matchsticks, to phase out production within three years and allot it to continue in labour-intensive cottage and small-scale units. The Minister pointed out that nearly 70% of India's matchsticks were being made by cottage industries, which employed 450,000 people. But the Western India Match Company, a subsidiary of a Swedish company, which makes the remaining

30%, was employing only 15,000 workers. Mr Charan Singh, who was Finance Minister in Mr Desai's government and is a fervent advocate of rural cottage industries, tried to force WIMCO out of the market by increasing excise duties on all machine-made matches and halving that on hand-made matches.

In the textile industry too, the Janata government pushed the existing policy to support handlooms further, by forbidding further expansion in the weaving capacity of textile mills. About ten million people are already employed in the handloom textile sector, which makes it the largest employer in India outside agriculture. But unfortunately most of them are only part-time workers and their wages are abysmal. Mr Fernandes tried to introduce a new technology to upgrade this industry and thus increase the productivity of handloom weavers and their wages. But he met with strong resistance from his own party. The Industry Minister tried to promote a technological breakthrough that allowed cotton-blended man-made fibres to be both hand-spun and hand-woven for the first time. The breakthrough had been achieved by a private Indian company which manufactures polyester fibres. The company estimated that to increase the availability of cloth in India by one metre per head by producing the new polyester-cotton handloom fabric, jobs would have to be found for at least 1.2 million people. But to make the new fabric competitive with hand-woven pure cotton, the government would have to drop the heavy excise duty that it currently levies on all man-made fabrics. As the polyester-blended fibre has a higher tensile strength than the pure cotton fibre, the high breakage rate would be reduced and the weaver's productivity and, thus, his wages would increase by about 30%.

But when Mr Fernandes tried to get the parliament to pass a bill that would allow the Khadi and Village Industries Commission to promote the new fabric, he ran into strong opposition. Several MPs suspected that the scheme would help the large polyester fibre manufacturers to escape excise duties and greatly expand their operations, while the public exchequer would suffer a great loss. Some

people also feared that the new fabric could drive those traditional weavers out of jobs who were not able to organise themselves sufficiently to benefit from the new technology. The new fabric could result in more unemployment. Mr Fernandes failed to get parliament's support.

However, an official in the Planning Commission who has long been a lone advocate of AT in official circles, pointed out that cement, for instance, even if it is produced on a mini-cement plant, will remain an inappropriate product for building houses in India as it is beyond the purchasing power of the majority of the country's population. Only a policy that promotes housing technologies which reduce the use of cement and increase the use of traditional building materials such as mud can be described as a policy for AT in the housing sector. But big ostentatious houses, built with a wasteful use of cement, are the order of the day amongst the Indian elite and such an approach to AT would require a very strong political will, something that the Janata government failed to display.

Meanwhile, however, Indian planners are certainly learning much to their chagrin that domestic technological capability can be a double-edged weapon. It can be used to subvert national objectives, just as much it can be used to serve them. The Planning Commission in 1977 had refused the State Government of Punjab permission to import combine harvesters. It had argued against these machines on the grounds of massive unemployment within the country. But the Punjab had created such a huge demand for labour that farmers within the region face crippling shortages of labour, particularly during periods of peak labour requirements. After the central government turned it down, the Punjab government approached the Punjab Tractors Limited (PTL) in Chandigarh, which it had helped to set up, to manufacture these combine harvesters. This company had been set up by an extremely enterprising group of former R&D engineers from the Central Mechanical Engineering Research Institute (a unit of the CSIR) to manufacture a small tractor that they themselves had designed. Production of the tractor started after considerable delay and opposition from those interest groups which wanted to collaborate with foreign companies. The efforts of this group for technological self-reliance have thus been widely praised and discussed. But now the PTL management cannot refuse to design and develop combine harvesters for the government of the state in which it is based even though it may not be happy to do so. The first prototype of an indigenously designed combine harvester is expected to be ready soon. Domestic technological capability thus would have served the Punjab government's purpose but subverted the Planning Commission's objectives. Now how does one control such a situation? □



Cement from rice-husk ash: but for whom?

UNCSTD: now we need action

Arnoldo K Ventura, Technical Director of the Scientific Research Council of Jamaica, counts UNCSTD 'a stitch in time'

After five preparatory conferences, several national, regional, and specialised meetings, UNCSTD begrudgingly approved a programme intended to tackle worldwide underdevelopment. But, because of the complexities of over-development and under-development, and the fact that the precise relationship of science and technology to development is still being unraveled, the conference was anchored on presumptions based on superficial observations and good intentions. Consequently, the resulting programme remains essentially a series of imprecise generalities which now beg for more precision and implementation.

The *ad hoc* mixture of politicians, diplomats, bureaucrats, and a few scientists, many of whom were attending the 'circuit' for the first time, or who were vaguely familiar with the nuances of the Conference, did not help the UNCSTD deliberations, while many of the few oldtimers in the UNCSTD extravaganza had already formed hardened opinions. Consequently, in spite of the fact that no legal codes were being considered, the debates were often long and tedious.

The debates lacked specificity, and it was clear that neither the developed nor the developing countries had a firm grasp on the systems, steps, or components of science and technology which are necessary to redress the problems of mal-development. Very little note was taken of the fact that while many studies have shown a strong relationship between rising industrialisation and the ferment of scientific technology, few categorically outlined comprehensive strategies for the resolution of underdevelopment based on the unique experiences of Europe, US, Japan or recently industrialised states.

As a matter of fact, the pivotal components of the science and technology process which have lead to the successes of the developed countries are still being determined. Consequently, the UNCSTD debate was replete with emotion and defensive rhetoric, rather than with compromise flowing from a framework of scientific facts or well-documented experiences. The "programme of action" used as the base document for negotiations was not action-oriented, as a result of the paucity of concrete information regarding the proper way to deploy science and technology to promote a balanced development within under-developed countries.

Although it is generally accepted that a decent level of indigenous technological competence is vitally necessary for this process to take place, the basic infrastructure necessary to provide this capacity in light of different concepts of development and differing national resource endowments, is still not fully

understood. Furthermore with limited technological capacities in many developing countries and with many countries still not sure of the type of investment or sacrifice needed to attain increased capabilities, the developed countries during UNCSTD tended to view the problem of underdevelopment in terms of what can, instead of what should, be done.

The ill-effects of some modern technologies on society, on culture, and on the environment have further confused the question of science and technology for development, thereby raising doubts about the transfer of the most modern technology for the amelioration of poverty. This type of confusion of the issues often lead to extreme positions, especially among delegates exposed to them for the first time. This engendered superfluous verbiage and a train of stalling manoeuvres. Nevertheless, in general, the debates have served to emphasise the fundamental nature of the problem and may bring solutions nearer.

Against this background, it was unfortunate that some specialised agencies decided unabashedly to use UNCSTD as an opportunity to expand their power base. This distraction from the main thrust of the Conference underscores the dilemma of poor countries, as often institutional politics looms to obliterate the pressing need to harmonize against the ravages of underdevelopment. This suggests the need for closer examination of the UN agencies and their contribution to the relief of human suffering.

Although much can be said about the imperfections of UNCSTD, its signal importance should not be overlooked. Apart from generating an increased awareness of the indispensability of science and technology to development among the developing world — the elaborate preparatory process for the Conference achieved that — the event itself brought together national decision makers and functionaries from the scientific communities to discuss concepts which are of vital interest to both. Such ventilation of the issues from different perspectives very seldom takes place at the local or national levels, and hopefully this may be the beginning of operational links between international political and the scientific circles, which may provide a clearer picture of the problems and bring us closer to collective solutions. UNCSTD gave the opportunity also for socially concerned scientists, from countries at different stages of development, to share experiences.

It is interesting to see how the voluntary contribution of some \$250 million for 1980 and 1981, earmarked for the promotion of science and technology, will be utilised in

the underdeveloped countries, granted the inadequate technological capacities in these areas, and the fact that the amount is paltry compared to the magnitude of the world's problems. It is clear that the major contributors will be tempted to control the expenditure of these funds, either to ensure that these allocations are used to establish a viable scientific and technological infrastructure or surreptitiously to ensure that these funds are not used to create undue competition for their local industries. It will be interesting to see how Third World scientists and technologists respond to the disbursement of these resources. UNCSTD, undoubtedly, will force some reaction from the scientific and technological communities of the Third World. Their response in turn will surely condition the attitude of their counterparts in the industrialised states in future.

There is also some hope for change — as the Conference has engendered more solidarity within the Group of 77, while there appear to be more areas of contradiction and differences arising among the ranks of the rich as they struggle to contain the poor countries in the status quo. Varying degrees of rivalry and contrasting views of the future have begun to prick the national consciences within the developed states. Further, the smaller and more susceptible rich countries find their interest merging closer to those of the underdeveloped states. This type of realism opens up avenues of salvation for mankind.

UNCSTD, if only for its size, has stirred the attention of the world. It has prompted new insights into the causes of the glaring disparities plaguing mankind. It has stimulated the closer examination of the relationship between science, technology, and societies, and it has bared fundamental questions concerning the future of mankind and the extent to which the global and local environment may be manipulated with impunity. New philosophies are surfacing as a consequence.

A realistic alternative to the arms race has been forwarded as scientists and technologists are being called upon to attack poverty instead of the human race. The concept of science being the common heritage of man is being extended to technology, which signifies the emergence of a new morality within the brotherhood of man. These are fundamental concepts which for expression will require basic changes in the psyche of the human race, and consequently, will be long in reaching maturity or practical fulfillment, but nevertheless they were being voiced by responsible delegates. Probably, UNCSTD is a 'stitch in time' which may save the international community from continuing along the path of global maldevelopment resulting from the misuse of the power of science and technology, and the espousing of bankrupt economic philosophies devoid of strong social probity. □

correspondence

Accounts of East European science are reliable

SIR, — Dr Lorch's attack on Vera Rich (13 September, page 98) is not only graceless; it is also futile. As many an obfuscating party bureaucrat knows to his chagrin, she is a tough and determined person who thrives on reckless abuse and has an uncanny facility for discovering awkward truths behind the glossy facades of official science. In other words, she brings to the readers of *Nature* the best-informed and most reliable account of science in Eastern Europe and the Soviet Union that is publicly available in the Free World. In so doing, Ms Rich complements the correspondingly unsycophantic reports of Colin Norman and David Dickson from Washington, and the sober dispatches that you have recently been publishing from various countries in the Third World. Good work; keep it up.

For it is not by enthusiasm for intergovernmental research projects, nor by the exchange of vacuous academic courtesies between scientists of rival nations, nor by cheerful splurges about marvellous work being done in this or that country, that the pursuit of science will be safeguarded. The fundamental factor for the health of the world scientific community is solidarity in protecting the human rights of scientists of all countries.

I believe that most of my colleagues with recent first-hand knowledge of the state of science in many countries of the Socialist bloc would agree that the situation is very largely as Ms Rich describes it, and that this continues to be a matter of the gravest concern for us all. The information that you thus make public in *Nature* is not only essential for those who would combat such evils; it is, by its very publication, of immense influence.

Yours faithfully,

JOHN ZIMAN

University of Bristol, UK

Doctor disappears in Afghanistan

SIR, — I am writing on behalf of the staff of this department, to bring to the attention of our scientific and medical colleagues, difficulties we are having in communicating with a former colleague in the Faculty of Medicine, University of Kabul, Afghanistan.

He is Dr Z. Razban, a medically qualified member of the staff of the Department of Medical Biochemistry in that university. He spent two years with us on a WHO Fellowship to gain experience in teaching biochemistry to medical students and also on research in immunological aspects of protein-calorie malnutrition. When he left us in January 1977, he was well trained in a variety of immunological techniques and had been awarded an MSc based on his research on "Acute phase proteins in children with protein calorie malnutrition" which was also published under this title (Razban *et al.* *J. Trop. Med. & Hyg.* 78, 264-266 (1975)).

At that time we understood he was the only doctor in Afghanistan trained in immunology, and on his return he intended to set up an immunological service and to train other doctors and technicians in these techniques, as well as continuing his research and teaching. In order for him to do this successfully it was

essential for him to maintain contact with us, and we find it disturbing that for reasons we are unable to discover, this contact has not been maintained. We are certain that Dr Razban would himself have continued to remain in contact with us, and can only assume that he is prevented from doing so by the authorities in Afghanistan.

Yours faithfully,

D.S. JACKSON

University of Manchester, UK

Tumour virus DNA: hazards no longer speculative

SIR, — The critical response (9 Aug, page 444) by Drs Rowe and Martin to the editorial (7 June, page 461) which called on the US Congress to enact recombinant DNA legislation, appears itself to be somewhat misleading.

While it may have been technically incorrect for the editorial to imply that it was unanticipated by tumour virologists that DNA from tumour viruses can cause tumours, it remains true that no one could have known prior to the Rowe-Martin risk assessment experiments (*Science* 203, 883 and 887; 1979; Israel *et al.*, *Science* in the press) that tumour virus DNA which has been cleaved with restriction enzymes and spliced into plasmids or lambda phage genomes would retain tumorigenicity.

The importance of this finding is the following: up until now opponents of strict recombinant DNA guidelines and legislation have portrayed fears that an aetiological agent could escape the laboratory in a novel biological form as "purely speculative." However, bacterial plasmids and lambda particles do not customarily carry agents of animal carcinogenesis into the ecological niches occupied by their bacterial hosts. It has now been shown by Rowe, Martin and their coworkers that it is a fairly routine matter to construct plasmids and lambda particles with just such a capability. The fact that these splice products are less effective than ordinary polyoma viruses in promoting tumours in the normal hosts of polyoma is irrelevant: the bacterial route for the dissemination and variation of tumorigenic DNA is a qualitatively new one.

The unlikelihood that the Rowe-Martin experiments released any novel tumorigenic vector into the environment is largely due to the fact that these studies were carried out under P4-EK2 containment protocols. However, under the newly revised NIH Guidelines (*Federal Register*, January 15, 1979) such tumour virus splicing experiments can now be carried out under P3-EK1 or P2-EK2 protocols.

It should be noted that the efficacies of both the biological safety cabinets used in P2 and P3 experiments, and of the only approved EK2 host, *E. coli* χ 1776, have been recently called into question (*Nature*, 29 March, page 384; S. Levy, technical report to NIH under risk assessment contract; R. Curtiss III, letter to NIH, May 11, 1979).

Since the hazards of releasing into the ecosystem known aetiological agents in new biological forms can no longer be considered speculative, it should now be unacceptable to denigrate efforts by the NIH Recombinant DNA Advisory Committee, and by the public through its unions and government, to tighten containment through guidelines and legislation

that extend to industrial as well as academic laboratories.

Yours faithfully,

STUART A. NEWMAN

New York Medical College, New York

Meeting on carcinogen testing is open

SIR, — In the 23 August issue (page 623) Alastair Hay wrote a short article "Press banned from conference" in which he discussed the status of press invitations to a meeting to evaluate short-term *in vitro* tests for carcinogenicity. The meeting is open to the public as are all such US Government sponsored meetings, and contrary to his headline and article the press has been neither banned, excluded nor specifically invited.

Investigators of the international programme for the evaluation of short-term tests for carcinogenicity will meet in October to present and discuss their results for testing 42 coded chemical samples. Results from a total of 68 assay system investigators will be presented. The meeting will not be comprised of formal presentations but rather of concurrent work groups. During the first part of the meeting five separate work groups will meet to discuss the raw data from similar bioassays. During the second part of the meeting the participants will be grouped into work groups to evaluate the cumulative test data on individual chemicals. Both phases of the meeting will require tedious analysis of large volumes of raw data and review of details of experimental protocols, neither of which is usually of interest either to the scientific community at large or to the press.

The coordinating committee fully recognise the importance of this study as well as the preparation of a final report for general distribution. It was this recognition that led to the decision to use two different mechanisms to bring this final report to the attention of the scientific and regulatory communities and the public.

Firstly a full account of the study will be published including the section on chemicals reviewed and carcinogenicity data on each chemical, results from individual investigators and an overall summary and evaluation. Secondly, a public information meeting has been planned for 3 December at the National Institutes for Health, Bethesda, to present and discuss the complete results of this study. It is the feeling of the coordinating committee that the six weeks between the October and December meetings will be necessary to make a comprehensive and meaningful evaluation.

It is regrettable that there has been such a misunderstanding concerning press participation at the two meetings planned to bring the international trial to completion.

Yours faithfully,

THE COORDINATING COMMITTEE

B A BRIDGES J ASHBY
P BROOKES
I F H PURCHASE T SUGIMURA

● *Alastair Hay writes: My experience, and that of other British science journalists, was that we applied to attend the meeting — only to be told that the presence of journalists might inhibit balanced scientific discussion. I subsequently had a letter from one of the organisers, confirming that no journalist was to be invited. If I misunderstood the committee's real intentions, then I apologise.*

news and views

Arctic animals and climatic changes

from Robert M. May

THERE is a tendency in much of the ecological literature to think of environmental change as taking place either on a very long, or a very short, time scale. Thus attention is given to the way populations, or even entire communities, respond to the major changes in climate that occur on geological time scales. At the other extreme, many studies concentrate on the influence short, year-to-year climatic fluctuations can have on the dynamics of populations (for example, in fish recruitment or locust outbreak). In a recent working paper of the Interim Committee on Marine Mammals of the International Union for the Conservation of Nature (ICMM/1/WP1), Laws, Gordon Clark and Dunbar note that there appear to be significant changes in climate on many intermediate time scales. They emphasise that such systematic environmental changes, particularly in the time range from around a decade to a century, raise important questions for conservation and resource management; these questions have conventionally received little attention.

Most people are familiar with the long sequence of alternations between glacial and interglacial epochs of climate, with a period of about 100,000 years, that has characterised the past million years or so. Mid-latitude temperatures today are 5°–8°C warmer, and sea levels 80–100 metres higher, than those prevailing during extreme glacial stages (the most recent being the Würm or Wisconsin glacial maximum, about 18,000 years ago). The World Meteorological Organization and others have used a variety of palaeoclimatic indicators to show that, following recovery from the Wisconsin ice age, the global climate has exhibited systematic fluctuations (albeit within narrower limits than those characterising the ice-age epoch), with expansions and retreats of polar ice and mountain glaciers at intervals of about 2,000–3,000 years. Nested within this is finer-scale structure: stratigraphy and pollen from Swiss and other glaciers shows five cycles of glacial advance and retreat over the past 3,000 years, with temperature minima around 1400–1300 BC, 900–300 BC, AD 400–750, AD 1200–1300 and AD 1600–1850. The 'Little Ice Age' of 1600–1850 is a well-documented historical event, during which

temperatures in northern Europe were 1° or 2°C lower than today and stormy conditions prevailed in the North Atlantic. Conversely, some of the intervening climates have been better than today; during the 'Little Optimum' of the 10th, 11th and 12th centuries (with a peak around AD 1080–1180) a Viking colony flourished in Greenland and wine was produced in England.

Laws *et al.* point out that "Within the lifetime of people at present in their 60's there has been a climatic cycle that has affected the economy of a number of countries." This cycle began on an upswing to warmer conditions about 1915, reached a peak in the early 1940s, and returned to colder conditions over the past 35 years or so. This downswing is probably still continuing. Older people's memories of the golden summers of their youth have a measure of objective reality. During this 60-year interval, temperatures in the regions of the West Greenland current, in Baffin Bay, in the waters of Iceland, and around Spitzbergen have fluctuated over many degrees in the atmosphere, and some 2°C in the hydrosphere. In the Antarctic, all stations with long records show trends in mean annual temperature over the past 75 years (1900–1975), of total magnitude 0.3°C at South Georgia, 0.7°C at McMurdo, 1.1°C at Laurie Island in the South Orkneys, and about 1.9°C for the Antarctic Peninsula (British Antarctic Survey Annual Report, 1976–1977). In South Georgia there has been a steady 33% increase in rainfall since 1906, while in the Scotia Sea there have been marked changes in the circulation (BAS Annual Report, 1975–1976). These temperature changes of a degree or two may sound modest, but they represent the transport of enormous quantities of heat.

Many of the grander patterns of human history can be reinterpreted in terms of the systematic climatic fluctuations of the past 3,000 years (see, for example, Ladurie, *Times of Feast, Times of Famine*, Allen and Unwin, 1976). There have undoubtedly also been marked effects on many animal populations, especially in the sea. Unfortunately, little direct evidence is available. One exception is the demonstration by Thoradotir (in *Polar Oceans*, ed. Dunbar, Arctic Institute of North America, 1977) that the primary

production north of Iceland decreased greatly by a factor of at least 2 or 3, between about 1965 and 1975, presumably in response to a cooling of the water and the southward extension of the limit of sea ice. Another important study is by Vibe (*Meddr Grønland* 170, 1; 1967), who uses bone material uncovered by archaeologists in their excavations of Eskimo and Norse settlements in Greenland, and accounts of sales to Greenlanders of rifles and other equipment, to get an idea of the way climatic fluctuations have influenced the population density and geographical distribution of animals in and around Greenland. Vibe considers ringed seals, harp seals, narwhals and white whales, polar bears, walrus, Greenland whale (or bowhead, *Balaena mysticetus*), blue and white arctic fox, ptarmigan, eider, reindeer, musk ox and cod. He broadly distinguishes three periods since the early 19th century. In the "drift-ice stagnation stage" (1810–1960) the climate of northern West Greenland is relatively cold, dry and stable. This climate is favourable for reindeer in central and northern West Greenland, and the population increases; populations of sea mammals and sea birds concentrate at central West Greenland. The drift-ice is also relatively stable throughout the winter off North-East Greenland, where the climate is favourable for the reindeer and the musk ox, whose populations flourish. In the "drift-ice pulsation stage" (1860–1910) the ice of the Arctic Ocean drifts into the Atlantic in larger amounts than before, and the climate becomes warmer with much precipitation and with great fluctuations. The populations of sea mammals and sea birds decrease in central West Greenland, and the wet winters are unfavourable for reindeer and musk ox, whose populations decline. In the "drift-ice melting stage" (1910–1960) the drift-ice of the East Greenland Current decreases, and the climate is relatively warm with some precipitation. Cod multiply in Greenlandic waters, and populations of sea mammals and sea birds increase in northern West Greenland and in East Greenland. The reindeer populations of West Greenland have ample summer grazing, but the winter

Robert M. May is Class of 1877 Professor of Zoology at Princeton University.

pastures are often covered by snow and ice, so the population fluctuates. Of particular interest is Vibe's comment that, "Until recently, like all other Arctic animals the Greenland Whale was very numerous during certain periods and less numerous during others. Its distribution and mass occurrence was determined more by movements in the drift-ice than by persecution from Eskimo and European whalers."

These climatic cycles, and the associated changes in the distribution and abundance of animals, have had major effects on the economies of Greenland and Iceland. The gist of Vibe's analysis is that, over the past 60 years, in West Greenland the whole ecosystem moved north with the warming and south again with the cooling, so that a seal-hunting economy was replaced by a cod-fishing economy, the latter now being in some jeopardy. Off Iceland, the herring fishery disappeared northward, and later reappeared in northern coastal waters. Similar events, also involving the Atlantic

cod, occurred around Spitzbergen. Discussing long-term trends in the climate, Cushing (*Science and the Fisheries*, Arnold, 1977) points to the alternation between the Norwegian herring fishery and that off the Bohuslan coast of Sweden, which has persisted since the Middle Ages, as being associated with the well-documented 100-year periodicity in the wind system of the northern hemisphere. There may be a similar 100-year periodicity in the fishery for the Japanese sardine.

In short, many different lines of evidence and inference suggest that systematic changes in climate take place on different time scales, and with various amplitudes. Laws *et al.* are clearly correct in arguing that "The shorter cycles, terms of a decade and up to half a century cannot be reasonably neglected for purposes of resource management." Two intertwined problems emerge: to achieve skill in forecasting such climatic changes; and to convince resources managers to include these factors in their deliberations. □

First discovery of fossil tree shrews

from R.D. Martin

THE recent identification in Indian Sivalik Miocene deposits of fossil material exhibiting clear affinities with modern tree shrews reported by Kaul *et al.* and Vasishat and Chopra in this issue of *Nature* (pages 213 and 214) fills a major gap in our understanding of this unusual group of mammals. The tree shrews achieved prominence some 50 years ago when Le Gros Clark pointed out in a series of papers a number of characters (notably in the morphology of the skull, brain and reproductive system) suggestive of an evolutionary relationship to the order Primates (*Proc. Zool. Soc., Lond.* 44, 461: 1924; *ibid* 44, 1053: 1924; *ibid* 45, 559: 1925; *ibid* 46, 1179: 1926). The transfer of the tree shrews (family Tupaiidae) from the order Insectivora to the order Primates was given an international seal of approval by Simpson in his influential classification of the mammals in 1945 (*Bull. Am. Mus. Nat. Hist.* 85, 1: 1945), and numerous papers followed on the supposed primate-like features of these semi-arboreal, South-East Asian mammals. However, since the 1960s a growing number of investigators have rejected the concept of a specific phylogenetic relationship between tree shrews and primates, and one of the most recent, authoritative primate classifications excludes the tree shrews altogether (Simons *Primate Evolution: An Introduction to Man's Place in Nature*, Macmillan, New York; 1972).

It is probably true to say that the modern consensus is that tree shrews are not related to primates after all. This view is partly based on additional evidence regarding cranial, dental and cerebral morphology along with reproductive characteristics and

biochemical aspects. However, reassessment of the methodology used for reconstructing phylogenetic relationships has had a far more fundamental and far-reaching influence on the elucidation of primate evolution. It is, in fact, still true that tree shrews share many striking similarities with primates. But it is also true that all of these similarities can probably be ascribed to one of the following two categories: (1) retention of characters unchanged from the common ancestral placental mammal stock; (2) convergent development of adaptations for life in the trees. A convincing case could only be made for regarding the tree shrews as primates if it were possible to postulate the development of specific characters in a definable ancestral stock derived from the ancestral placentals and giving rise to tree shrews and 'other' primates. Exclusion of the tree shrews from any phylogenetic link with the primates owes as much to increasing attention to the principles of phylogenetic reconstruction as to the discovery of new evidence. As Romer suggested (*Notes and Comments on Vertebrate Paleontology*, University of Chicago Press, 1968), it now seems likely that the tree shrews, rather than providing a living model for the ancestral primate (the view espoused by Le Gros Clark and Simpson), probably provide a better indication of the characteristics of the common ancestor of all living placental mammals.

One major drawback in discussions of the phylogenetic ties of the tree shrews has always been the absence of any clearly identifiable fossil relatives. The Oligocene *Anagale* was, for a while, regarded as an

early relative of the tree shrews (Le Gros Clark *The Antecedents of Man*, 2nd ed., Edinburgh University Press, 1971), but has now been assigned to the Lagomorpha (rabbits and hares) (McKenna *Am. Mus. Novit.* 2158, 1: 1963). Van Valen (*Evolution* 19, 137: 1965) suggested *Adapisoriculus* as a Late Palaeocene tupaiid, but the fact that this fossil genus has been variously assigned to the marsupials, the insectivores and the tree shrews simply underlines the very generalised morphology of *Adapisoriculus* and the modern Tupaiidae. Strangely, although inferred phylogenetic relationships among mammals have been traditionally based heavily on dental morphology, the dental characters of tree shrews have been barely mentioned in discussions of their evolution. In fact, they exhibit a combination of somewhat unusual specialisations of the anterior teeth with a basically primitive configuration of the cheek teeth. Apart from the loss of one premolar in both upper and lower jaws, they depart only slightly from the pattern postulated for ancestral placental mammals.

The new material from the Sivaliks assigned to *Palaeotupaia sivalicus*, in spite of its fragmentary nature and the absence of any information on the lower jaw, shows clear similarities to modern tree shrews (actually to the Tupaiinae, rather than to the Ptilocercinae). The upper cheek teeth have a similar, basically primitive configuration, while the anterior part of the maxilla together with a premaxilla is similarly drawn out into a rostrum with a small canine posteriorly and two large, well-spaced incisors arranged longitudinally and separated by a large median gap from their counterparts. The upper dentition apparently had the same formula (2.1.3.3) as in all modern tree shrews, and the overall arrangement of the dentition, upper jaw, snout and orbital region is remarkably like that of modern *Tupaia*. Interestingly, although the fragments indicate a skull size comparable to that of the most arboreal modern *Tupaia* species (*T. minor*), the snout is longer and generally more like that of the more typical semiterrestrial *Tupaia* species. This confirms the suggestion (Martin *Man*, 3, 377: 1968; *Symp. Zool. Soc. Lond.* 33, 301: 1973) that some of the primate-like features of modern *Tupaia minor* (Le Gros Clark *The Antecedents of Man*), not seen in semiterrestrial or terrestrial tree shrew species, are independent acquisitions associated with increased arboreal adaptation and as such are of no direct relevance to primate ancestry.

Unfortunately, we still have no fossil evidence relating to the early origins of the tree shrews or their evolution over some 50 million years or more up to the Miocene,

R.D. Martin is Reader in Physical Anthropology at University College London

but the new Sivalik finds do at least document the presence of tree shrews of modern aspect in India some 10 million years or so ago, perhaps contemporaneously with such undoubted prosimian primates from the Sivaliks as *Indraloris* (Lewis *Am. J. Sci.* **26** 134; 1933; Tattersall *Postilla* (Yale) **123**, 1; 1968). □

Origin of the Galilean satellites and Jovian rings

from David W. Hughes

REGULAR satellite and ring systems seem to be a common feature of the major planets. These systems are coplanar, the orbits are circular and also the ratio between the orbital radius of a satellite and that of the one immediately interior is a constant, about 1.65 in the case of Jupiter.

The formation of satellites as the planets condensed has often been linked with the formation of the planets as the Sun condensed, and this process has been re-investigated by Prentice and ter Haar in a recent edition of *Nature* (**280**, 301; 1979). These authors return to the hypothesis originally advanced by Laplace in 1796. In the beginning, according to Laplace, the solar and planetary material was in the form of a huge spinning gaseous globe with a diameter larger than that of Neptune's orbit. As this great globe cooled it contracted, spinning faster and faster due to the conservation of angular momentum. A stage was eventually reached when the equatorial centrifugal force became larger than the gravitational force and a ring of matter was left behind, suspended in a circular orbit in the equatorial plane. Due to this mass and angular momentum loss, the proto-Sun could contract a little further, but it soon became unstable again and shed another ring. This process was repeated until the Sun reached its present size.

Now it must be remembered that the present mass of the Solar System is only a very small fraction of the mass of the rings, simply because most of the planets consist of material that is quite rare by normal cosmic standards. Each planet is probably the condensation remnant of a ring of about a thousand times the Earth's mass. Also the spinning gaseous globe was probably strongly turbulent. By considering T. Tauri stars Prentice (*Astr. Astrophys.* **27**, 237; 1973) concluded that the convective velocities in the globe are supersonic and that the turbulent stresses set up by these convective movements are of the order of 30–100 times the normal gas pressure just below the surface. Two things result from this. First, the globe becomes very centrally condensed which obviously leads to a low moment of inertia. Second, a dense shell of non-turbulent gas develops at the surface, this evolving into an

equatorial belt as the cloud contracts. It is this belt which is left behind when centrifugal force exceeds gravitational force, and it can be shown that the ring-shedding process repeats itself at a sequence of orbital radii R_n which satisfy

$$R_n/R_{n+1} = (1 + m/M_e f)^2.$$

In this equation m is the mass of the ring, M_e is the mass of the gaseous globe that has just lost the ring and f equals $I_e/M_e R_e^2$ where I_e and R_e are the moment of inertia and radius of the globe.

For the specific case of the Galilean satellites of Jupiter Prentice and ter Haar set M_e to 1.9×10^{30} g the present mass of Jupiter, f to 0.02 and R_n/R_{n+1} to 1.65. This immediately tells us that the mass of each Jovian ring was about 1.1×10^{28} g. If this mass was of solar composition it would have two medium-temperature ($\sim 0^\circ\text{C}$) solid components. The rocky one would have a mass of about 5×10^{25} g and the water ice and rock mixture component would have a mass of about 12×10^{25} g. The globe is warming up as it condenses so each detached ring will have a different temperature. The authors calculate that the temperatures at detachment for Io, Europa, Ganymede and Callisto are 533, 338, 220 and 133 K respectively. It is this temperature difference which causes the chemical variation. Io and Europa with masses of 9×10^{25} and 5×10^{25} g, respectively, have little ice, the high-temperature Io condensation having iron in the form of FeS while for Europa, Fe appears as FeO. The heat released as the ring condenses to form a satellite and as the satellite compacts causes S and H₂O to leak out, leaving sulphur deposits on Io and icy plains on Europa. Ganymede and Callisto condensed at a temperature below the freezing point of water and therefore consist of dirty ice.

By extending this theory, the authors predict that the contracting gaseous globe can shed one or two more rings interior to Io. One has a radius of $2.5 R_p$ and the other $4 R_p$ where R_p is the present Jovian polar radius, 6.677×10^9 cm. Amalthea is near the first of these but is spectrally very close to a C-type asteroid and has probably been captured by Jupiter at a later date. These interior gaseous rings probably did not condense to form satellites because there was too little time for the rocks to aggregate into a single body before the gaseous component of the rings dispersed, taking with it the finer dust material. Also the increasing luminosity of proto-Jupiter probably frustrated the settling out process.

Ring shedding cannot proceed indefinitely. The rate of release of gravitational potential energy slows down between $3 R_p$ and $1.5 R_p$ in the case of Jupiter. This results in a decrease in the surface temperature and in the supersonic

turbulence. R_n/R_{n+1} drops to 1.02, leaving the rings so closely spaced that they essentially merge together to form a continuous disk. The disk does not coalesce into a few single bodies because of its differential rotation. The gaseous component disperses and orbital resonance subsequently changes its structure. It is these resonances which may account for the sharp ring edge seen at $1.9 R_p$ by the Voyager I and II spacecraft.

No theory in cosmogony is without its rival and the origin of the Jovian rings is no exception. Smoluchowski wrote in the following issue of *Nature* (**280**, 377; 1979) that the rings were probably produced by the tidally induced fragmentation of a non-spherical, mechanically weak carbonaceous body. This disruption is thought to have occurred fairly recently. The solid surface area of the rings has been estimated to be about 20,000 km², this value coming from a comparison between the decrease produced in the charged particle flux by the rings and that caused by Amalthea, both these effects being observed by Pioneer 11. Bombardment by the intense flux of protons in the Jovian Van Allen belts erodes particles. Smoluchowski calculated that ice was eroded at between 50 and 100 Å yr⁻¹ and silicate particles at 2–3 Å yr⁻¹. Many small grains are thus eliminated on the time scale of a million years and the Jovian rings are not expected to have a lifetime much longer than this. □

Normal fluid density of superfluid ³He-B

from P. V. E. McClintock

NEW high precision measurements of the normal fluid density of liquid ³He-B have yielded some surprises. In particular, it seems that the so-called 'stripped' normal fluid density is independent of pressure, quite contrary to expectation. The experiments were carried out by Archie, Alvesalo, Reppy and Richardson of Cornell University and are reported in a recent issue of *Physical Review Letters* (**43**, 139; 1979).

The hydrodynamic properties of the B-phase of ³He are relatively simple in that the liquid can conveniently be described in terms of a two-fluid model, qualitatively very similar to that which has been used in relation to superfluid ⁴He for many years. Thus, the liquid is envisaged as being composed of two quite separate but completely interpenetrating fluids: a normal fluid component whose properties are similar to those of conventional liquids; and a superfluid component able to indulge in dissipation-free flow, carrying no entropy, and responsible for much of the seemingly bizarre behaviour exhibited by

David W. Hughes is a Lecturer in the Department of Physics, University of Sheffield.

the liquid. Properly speaking, of course, because ^3He is a fermion (containing an odd number of fundamental particles) its superfluid phases are more closely analogous to the gas of electrons (also fermions) in a superconducting metal than they are to superfluid ^4He , which is a boson (even number of fundamental particles). Thus, the origin of the two-fluid behaviour can be understood in terms of a modified form of the well known BCS theory of superconductivity, with an energy gap separating the ^3He atoms which exist in the paired zero-energy state (superfluid component) from the atoms in ordinary unpaired states (normal fluid component). In the A-phase, the energy gap is anisotropic, that is, the effective magnitude of the gap depends on direction within the liquid. It follows that the normal fluid density also depends on direction, rendering the hydrodynamic properties complicated in the extreme. The energy gap and normal fluid density in the B-phase, on the other hand, both turn out to be isotropic, and so the situation is a good deal simpler.

Measurement of the normal fluid density ρ_n as a function of T/T_c is potentially able to provide a great deal of information about the attractive interaction between the ^3He atoms which gives rise to the formation of pairs, and which is therefore responsible for the superfluidity. Here T is the temperature and T_c is that of the superfluid transition. The interaction is in fact extremely complicated, because the very act of forming the pairs in turn modifies the interaction. Much effort is being directed towards gaining an understanding of this so-called strong coupling effect.

The technique used at Cornell for measuring ρ_n entails shaking the liquid back and forth in a container sufficiently narrow that viscous forces constrain the normal fluid to move, thus contributing to the moment of inertia, whereas the superfluid component slips freely and makes no contribution. The apparatus involves liquid ^3He held inside a cell shaped like an inverted mushroom, with a thin coin-shaped cavity in its cap. The stem of the mushroom is a springy beryllium-copper tube which is used not only for filling the sample of ^3He but also for providing the restoring force enabling the cap to undergo torsional oscillations. From very careful measurements of the resonant frequency of such oscillations it is possible to deduce the proportion of the liquid which follows the motion of the cap, and hence to calculate the normal fluid density. This type of apparatus has been used before at Cornell. The main importance of the present results is their comprehensiveness, particularly in covering so wide a range of temperatures, achieved by incorporating the experimental cell into a nuclear demagnetisation cryostat.

P.V.E. McClintock is a Senior Lecturer in the Department of Physics at the University of Lancaster.

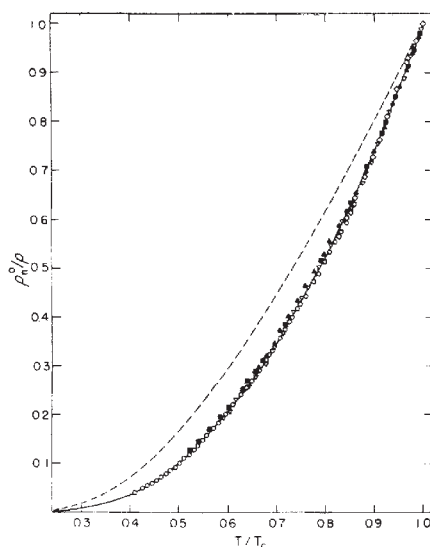


Fig.1 The stripped normal fluid density ρ_n^0/ρ (see text) plotted against T/T_c , where T is the temperature and T_c that of the superfluid transition. On this type of plot the experimental data, recorded for five different pressures between 2 and 29 bar, are clearly pressure-independent and lie well below the weak-field BCS prediction (dashed curve).

Values of ρ_n/ρ were measured as functions of T/T_c for several pressures between 2 and 29 bar. Here, ρ is the total density of the liquid. For any given value of T/T_c , it was found that ρ_n/ρ decreased with increasing pressure. The real interest arises, however, when the data are manipulated to give what the authors describe as the "stripped normal fluid fraction", ρ_n^0/ρ ; in effect, this amounts to removing the known effect of the Fermi-liquid molecular field, the interaction which is present in the normal (non superfluid) liquid at temperatures above T_c . Any deviations between ρ_n^0/ρ and predictions based on the BCS approach can then be attributed to strong coupling effects. The result of this operation is shown in Fig. 1. It is immediately apparent that, within the (remarkably small) random experimental error, ρ_n^0/ρ is completely independent of pressure. Also, in that there is a sizable difference between the data and the BCS theoretical curve, it appears that there are substantial strong coupling effects, as expected. It would seem, therefore, that the strong coupling effects are themselves independent of pressure.

This is a most unexpected conclusion. It seems to be at variance with other measurements of the strong coupling corrections based on specific heat or sound attenuation data, but to be in reasonable agreement with measurements based on other types of hydrodynamic experiment, once these have been readjusted in the light of the latest temperature scale. Devising an explanation of these results will be an interesting challenge for theorists working in the area and should eventually lead to an enhanced understanding of both superfluid phases of ^3He . □

All sorts of protoplasts

from Edward C. Cocking

ONE of the most refreshing aspects of protoplast research is that there always seems more to do. At the recent 5th International Symposium*, which covered most aspects of plant, fungal and bacterial protoplasts, it was clear that many laboratories are now using this novel cell system for many basic and applied investigations. Some of the basic studies discussed included pioneering work by L. Vigh and his colleagues at the Biological Research Centre, Szeged on the incorporation of spin-labelled fatty acids into the plasma membranes of intact wheat protoplasts. Significant differences were found between the microviscosity of membranes of cold-resistant and cold-sensitive wheat cultivars indicating that membrane fluidity may be directly involved in avoiding freezing injury — perhaps growth substance workers could find this a useful probe for early effects on the growth of single cells.

With the current interest in genetic manipulation, it was understandable that there should be a major emphasis on the fusion of protoplasts for somatic hybridisation, and on the use of protoplasts for transformation. A particularly useful feature of this meeting was that it afforded the opportunity of comparing progress between prokaryotic and eukaryotic systems. Classical genetic information transfer in bacteria is unidirectional, contrasting with the products of induced protoplast fusion which produces a new entity containing the whole genome and cytoplasm of two or more bacterial cells. L. Alföldi and his colleagues (Biological Research Centre, Szeged) highlighted the limitations of classical methods of genetic analysis when applied to such bacterial protoplast systems. The industrial potential of these interesting recombinants has not, apart from work on *Streptomyces*, been fully exploited yet.

However, the outlook for fungal protoplasts is more encouraging. Many workers have investigated the fusion of the protoplasts from a wide range of fungi. New genetic variability is being produced as a consequence of fungal protoplast fusion, particularly between various *Candida* species that are important in single cell protein production. These basic studies are providing a foundation for many other industrial applications (see Peberdy *Proc. 3rd Int. Symp. Gen. of Industrial Microorganisms*, 1978). Better strains of *Trichoderma* for improved cellulase production are being developed and brewing yeasts could benefit from the readiness with which polyploid strains can

*This meeting on Microbial and Plant Protoplasts was held in Szeged, 9-14 July, 1979, jointly organised by the Department of Microbiology, Attila Jozsef University, and the Institutes of Genetics and Plant Physiology, Biological Research Centre. The proceedings will be published by Pergamon.

be produced by protoplast fusion methodology. These studies and a refinement resulting in interspecific mitochondrial transfer (L. Ferenczy, Atila Jozsef University, Szeged; R. Dales & J.H. Croft, University of Birmingham, Birmingham), and the now well-documented transformation of yeast protoplasts by chimaeric plasmids, have fully established fungal protoplasts in the forefront of genetic manipulation.

Studies on the reversion of fungal protoplasts were however, not neglected (O. Necas, J.E. Purkyne University, Brno) and the report on chitosomes and the origin of chitin microfibrils (Bartnicki-Garcia, University of California, Riverside) showed the usefulness of protoplasts for elucidating the details, both structural and biochemical, of cell wall synthesis at a naked plasma membrane surface.

Close parallels exist with plant protoplasts in this respect, but the more exacting developmental processes involved in whole plant regeneration have restricted the range of species for plant genetic manipulations. Several workers emphasised the difficulty in releasing the probable inherent totipotency of single cell cereal cultures—isolated cereal protoplasts are the 'acid test' for such regeneration capability. Chinese workers reported limited success in obtaining a few cell divisions in cultures of wheat leaf mesophyll protoplasts. Protoplasts isolated from suspension cultures derived from callus of pearl millet immature embryos were reported to divide to form embryoids and subsequently whole plants (I.K. Vasil & V. Vasil, University of Florida, Gainesville). Several participants doubted whether this system would be reproducible, and also whether the approach adopted would be applicable to the main cereal crops; but this report will no doubt greatly stimulate work on cereal protoplasts.

Virus replication studies using plant protoplasts do not depend on such a full developmental sequence; and have provided an important clue to the various resistance barriers that may be by-passed using protoplasts: I. Takebe (Nagoya University, Nagoya) reported that radish mesophyll protoplasts are readily infected by brome mosaic virus which shows no evidence of infection in radish leaves.

A somewhat similar breakdown of infectivity barriers is suggested by the reports of the transformation of plant protoplasts by isolated *Agrobacterium* plasmids—in future it may prove possible to transform protoplasts of monocotyledons although *Agrobacterium* itself infects dictotyledons only. If this is coupled with the ability to regenerate whole plants, a sound basis could be provided for the transformation of cereals.

Understandably most of the somatic hybridisation work using plant protoplasts

has utilised more culturally amenable species. Although several somatic hybrid plants have been obtained between species that have not, as yet, been crossed sexually this has not resulted in the production of useful amphidiploids. It is still not clear whether chromosome elimination in heterokaryons of disparate species will result in cybrid formation, with the possibility of mitochondrial recombination as suggested from the restriction nuclease analysis of G. Balliard and G. Pelletier

(University of Paris-Sud, Orsay) or regeneration of plants with only very limited gene transfer. There is obviously the possibility of such somatic transformation using disparate plant protoplast combinations (D. Dudits *et al.*, Biological Research Centre, Szeged).

There was general agreement that only continuing research on all these various aspects of protoplasts will provide an adequate foundation for applied developments. □

Twenty years of pheromones

by Mary Lindley

THE exchange of chemical signals between living organisms has proved to be far more complicated than it seemed twenty years ago, when pheromones were first defined as substances responsible for chemical communication among individuals of the same species. Through the combined efforts of chemists, ecologists and physiologists, concentrating chiefly on insects, the original idea that one substance served one species has given way to the realisation that the pheromone secreted by a particular species is a characteristic mixture of components. Each component has a particular function in communication within and between species, affecting the behaviour involved in, for example, attraction, inhibition and alarm. Electrophysiological studies of insect chemoreceptors, such as those on antennae, have shown that different cells respond to different components, and that the central nervous system interprets the resulting signal according to the level of excitation of the various cells.

The realisation that there is more to chemical communication than sexual attraction has necessitated a more extensive classification of the chemicals involved. When individuals of different species communicate chemically, the signal is a kairomone if the receiver benefits, as in the case of certain predators of bark beetles which aggregate in response to secretions from their prey. If the sender benefits, however, the signal is an allomone, such as those produced by orchids of the genus *Ophrys* to attract the bees that pollinate them. Some of the contributors to the growing understanding of chemical communication gathered recently in Sweden*. Among the inevitable and essential accounts of further pheromone components isolated and synthesised, there were some descriptions of newly recognised chemical relationships and progress reports on the application of pheromones in pest control.

As sex attractants, pheromones have been considered primarily to be produced by females to attract males, but this is not always the case, and D. Schneider (Max-Planck-Institut für Verhaltenphysiologie, Seewiesen) described one produced by

male monarch butterflies (family Danaidae). Through thin hairy organs, known as hairpencils, they release a pheromone to prepare the females for mating. Its principal component is a heterocyclic ketone—danaidone—synthesised from precursor alkaloids taken up by the butterflies from certain food plants. Synthesis of danaidone also often requires contact between the hairpencils and the male's wing gland. Together with cardiac glycosides, the alkaloids also serve as chemical defence agents which render the butterflies unpalatable to some of their predators.

R. Baker (University of Southampton) believes that some of the most successful West African predatory ants have become specialised so that their pheromones, used in alarm and recruitment of nest mates, are not detected by their prey (*Experientia* in the press). The workers of some species of myrmicine ants can move freely within the galleries of termites before attacking the inmates and carrying them to their own nests. But when the Southampton group exposed termites to workers of related species of ants, the termites retreated. Bioassays and analysis of the secretion of the mandibular gland in various myrmicine ants have shown that the specialised or semi-specialised predators produce aliphatic alcohols, which do not affect termites, instead of carbonyl or other compounds, which are usually repellant to termites. Baker suggested that the secretions of the ant's mandibular glands may provide clues by which termites recognise predators, and that the most successful predators are those producing secretions the termites cannot detect.

When attention turned to pest control, there were enthusiastic reports of some of the programmes that have been set up throughout the world. The general response, however, was less enthusiastic, with discouraging references to unobtainable patents and the strict standards of environmental agencies more

*A conference entitled 'Chemistry of Insects: Chemical Interaction and Communication' was held in Borgholm, Sweden on 13-17 August 1979. It was organised by G. Aulin-Erdtmann of the Swedish National Committee for Chemistry of the Royal Swedish Academy of Sciences, and sponsored by Euchem and the Swedish National Committee for Chemistry.

Edward C. Cocking is Professor of Botany at the University of Nottingham.

familiar with toxic pesticides than behaviour-modifying chemicals. Although nobody could foresee pheromones alone taking care of pest control, there was nevertheless a general hope that they would decrease the present dependence on pesticides.

Mass trapping — large scale capture of pests in traps baited with pheromones — forms the basis of several programmes. A. Shani (Ben-Gurion University of the Negev, Be'er Sheva) pointed out that Israeli farmers spray their cotton fields with pesticides less often than was necessary before mass trapping of the Egyptian cotton leaf worm, *Spodoptera littoralis*, began in 1975. Mass trapping is also the basis of the Norwegian campaign against the spruce bark beetle, *Ips typographus*. Although the beetle is usually a valuable member of the forest ecosystem, accelerating the decay of dead wood, its population has grown so rapidly during the past 10 years that much of Norway's spruce forest is threatened. L.A. Riege (Borregaard Industries Ltd, Sarpsborg) felt that there were signs of success in the catches of the 600,000 traps set up this year with a lure of bark beetle pheromone.

G.N. Lanier (State University of New York, Syracuse) reported progress in trials with the aggregation pheromone of the European elm bark beetle, *Scotylus multistriatus*, the carrier of the fungus that causes Dutch elm disease. Lanier has seen a decline in the incidence of diseased trees in trial areas during the past year. He spoke enthusiastically of the potential of this sort of control programme if accompanied by other precautions, such as the proper removal of infected wood.

L. Rechav (Rhodes University, Grahamstown) has been attacking the tick *Amblyomma hebraeum* by painting cattle with a mixture of tick pheromone and acaricide. Small areas of the mixture attract unfed nymphs and adults, killing them before they bite the cattle, thus reducing the chance that they will spread disease. However, this method has the disadvantage that tick assembly pheromones are species specific and remain effective on the cattle for only three days, precluding widespread application at present.

The aim is not always to trap and kill insects, but sometimes to confuse them by filling the atmosphere with pheromones so that the males cannot find the females. The only effective way to monitor the results of such methods is to count the larvae that hatch in a test area, and the results can be distorted by the arrival of gravid females from untreated areas. In tests of the efficacy of mating disruption by pheromone inhibitors of the striped rice stem borer, *Chilo suppressalis*, at the International Rice Research Institute in the Philippines, this problem has been overcome by the use of large field cages.

Mary Lindley is an Assistant Editor of Nature.

P.S. Beevor (Tropical Products Institute, London) said that it had been possible to quantify rates of release that would disrupt mating up to 100% in this way. The next stage will be to microencapsulate the inhibitors for slow continuous release and set up trials in the open field. As with similar test programmes, however, the outcome cannot be predicted.

A third practical application of pheromones is to attract pests for population monitoring. Baker reported the early stages of a programme to monitor the pine beauty moth, *Panolis flammea*, a pest of lodgepole pine in Scotland. Once the most effective formulation of pheromones has been worked out, traps will be set up to capture sufficient moths to reflect variations in the size of the populations.

Participants were glad to note that monitoring is already routine in many areas, with commercially produced traps, pheromones and dispensers being used to forecast the size of pest populations. On the basis of the numbers caught in the traps, farmers and their advisors decide how often to spray with pesticides, and there has been a general reduction in the frequency of spraying. Pests being monitored in this way include the pea moth, *Laspeyresia nigricana* in England and northern Europe, the codling moth, *L. pomonella* throughout Europe, *S. littoralis* in Israel and the pink bollworm,

Pectinophora gossypiella in California.

Pheromones may have uses beyond the continuing battle against pests. The declining population of honey bees, *Apis mellifera*, in Britain and the unattractive nature of arable crops have stimulated attempts to use pheromones to bring the two together for pollination of the crops. J.A. Pickett (Rothamsted Experimental Station, Harpenden) described how this work has brought an unexpected insight into the way the bee apparently controls its output of pheromone components. The Nasunov gland, between the sixth and seventh segments of the body, produces several components, of which (E)-citral is the most active in attracting other bees, although it is very volatile and present in very small quantities. Pickett and his colleagues have found that the gland has a highly specific enzyme system that converts geraniol to (E)-citral. He suggested that this system controls the release of (E)-citral, ensuring that it is present in sufficient quantity to be active.

The way in which insects synthesise their pheromone components and release them in precise proportions are problems for the next twenty years, together with the mode of action at the receptors and the subsequent fate of the pheromone. Meanwhile there are plenty more pheromones to be isolated from an abundance of insects. □



A hundred years ago

The French Association for the Progress of Science has taken a bold step and decided that its session for 1881 will take place in Algiers. To avoid the numerous inconveniences of the strong heat which prevails all over Algeria in the month of August, it has been decided that the meeting should be held in April, during the Easter recess.

This happy result of the deliberations must be attributed to the personal exertions of M. Albert Grévy, the brother of the President of the French Republic, who holds the post of Civil Governor-General of Algeria.

It is supposed that the new scientific establishment whose formation has been decreed this year, will be formally inaugurated on this occasion, and a scientific movement of some importance will take place in the colony.

The *Akhbar* announced a few days back that a geographical society is being organised in Algiers.

In the meantime a number of representatives headed by Algerian senators and deputies will make a tour of exploration during the month of October. They will start at the end of September, as we announced some weeks ago. They will witness an agricultural and horticultural exhibition, which is to be held at Bone, for the whole of Algeria and Tunisia, and which will be held in Algiers in 1881.

The most successful lecture this year at Montpellier, was organised by the Languedocian Society of Geography. MM Solcillet, Brau de Saint Pol Lias, Director of the Sumatra Exploring Company, and other explorers or intending explorers, appeared before the public on that occasion. M. Rabaut, the President of the Society of Geography of Marseilles, and the commercial agent for the Sultan of Zanzibar, gave most interesting details of the several explorations at present going on in that part of the Dark Continent.

The lecture on the progress of electricity was given by M. Denayrouze, of Jablochkoff candle notoriety. The speaker tried to show that Jamin's candle ought to be superior to the light which is spreading so largely in Paris and in London.

The British Association

Report of the Committee for exploring Caves in Borneo. Drawn up by Dr. J. Evans F.R.S. — The Committee report that with the grant of 50 l. from the Association, a similar grant from the Royal Society, and a further sum of 200 l. from private sources, they have been able to prosecute an examination of various caves in Borneo, under the superintendence of Mr. A.S. Everett, who has devoted himself to the task for a period of nearly nine months.

The final report upon his work has not yet been received, but it appears from his letters and from the specimens which have been transmitted to this country, that nothing of special interest either from an anthropological or a geological point of view has resulted from his explorations.

From *Nature* 20, 18 September, 484, 499; 1879.

articles

A binary star model for SS433

D. S. De Young

National Radio Astronomy Observatory, Charlottesville, Virginia 22901

Geoffrey Burbidge

Kitt Peak National Observatory, Tucson, Arizona 85726

The novel features recently observed for the strong emission line object SS433 can be explained by a model in which SS433 is a binary of a massive early-type star and a magnetic white dwarf.

RECENT observations by Margon *et al.*^{1,2} and Liebert *et al.*³ of the strong emission line object SS433 have revealed some extremely interesting and unusual features. Most prominent among them is the presence of strong emission lines of hydrogen and neutral helium which show simultaneous, large redshifts and blueshifts of up to $+50,000 \text{ km s}^{-1}$ and $-30,000 \text{ km s}^{-1}$. These shifts are relative to a nearly stationary line system¹ of velocity $\leq 180 \text{ km s}^{-1}$. The detailed line structures are complex and show smaller velocity variations on a scale of days^{1,3}, as well as variations of linewidth from 50 to $\sim 200 \text{ \AA}$. A second unusual characteristic is that the red- and blueshifted systems merge to a common velocity⁴ of $\approx +10,000 \text{ km s}^{-1}$ with an apparent period of $160 \pm 3 \text{ days}$ ^{2,4}, thence returning to high red- and blueshifted values. Moreover, there seems to be no evidence of eclipsing of the high velocity features during a period; the period is consistent with a sinusoidal form, and there may be evidence for non-Keplerian motion in the high velocity features⁴. Margon *et al.*¹ deduce a distance to SS433 of $\geq 3.5 \text{ kpc}$, giving $M_v \approx -3.5$. To a limit equivalent to that of the Sky Survey plates, the object appears stellar and is considerably reddened¹. A search for optical polarisation³ has produced an upper limit of 0.1%. Seaquist *et al.*⁵ have observed SS433 at radio frequencies over a period of 18 months and found variations on a scale of hours to weeks at centimetric wavelengths, with changes of flux of up to factors of ~ 2.5 (200 to 500 mJy) on the longer time scale. Neither linear nor circular polarisation was detected at 2.7 and 8.1 GHz, and spectral information available suggests a non-thermal origin for the radio emission. VLBI observations of SS433 at 21 cm show no structure on a scale of several milliarc seconds down to a limit of $\approx 20 \text{ mJy}$ (R. Sramek, personal communication). Finally, the object may be associated with the X-ray source A1909+04¹.

Constraints on models

A viable model for SS433 must primarily account for: (1) the large luminosity in the emission lines of $\sim 10^{34} \text{ erg s}^{-1}$; (2) the large redshifts and blueshifts relative to a nearly stationary system; (3) the periodic merging of these features to a value of $+10^4 \text{ km s}^{-1}$; (4) the stellar appearance of the object; and (5) the observed emission at radio and X-ray frequencies. For a brightness temperature of $\sim 10^4 \text{ K}$, (1) implies that the $\text{H}\alpha$ emitting region must have a size $\sim 10^{12} \text{ cm}$ if the emission comes from an optically thick gas, as appears to be the case³. This eliminates accretion onto a neutron star, since the accretion radius is much smaller than 10^{12} cm . Similarly, Zeeman splitting in the

magnetosphere of such a star to produce (2) is eliminated because no significant optical polarisation is present³. Terlevich and Pringle⁶ have proposed a model for SS433 which involves accretion onto a $3 \times 10^5 M_\odot$ black hole, while Fabian and Rees⁷ have suggested that oppositely directed high-speed jets emanate from a central source of unknown nature. Both of these models, in their present form, can account for the observed characteristics except for the periodic merging to a high velocity required by constraint (3).

A binary star model

We propose that SS433 is a member of a class of astrophysical systems whose existence has been confirmed by observations; that is, we suggest that SS433 is a binary star system comprised of a massive early-type star and magnetic white dwarf. As we describe below, the observed characteristics of the system can arise if the rotating white dwarf possesses gas imbedded in a dipole field either aligned with or at large obliquity to its spin axis, and with the plane of the orbit lying nearly in the plane of the sky. Precession, but more probably nutation, induced by dissipative interaction with the primary can give rise to the apparent periodicity.

Production of the shifted emission lines: The general characteristics of magnetic white dwarf stars have recently been reviewed by Angel⁸. Typical masses and radii are $\sim 1 M_\odot$ and 10^9 cm . Inferred surface magnetic fields range from 10^7 G to $> 10^8 \text{ G}$; the upper limit of the surface field is uncertain, and 10^9 G is not excluded. A firm upper limit is reached at 10^{12} – 10^{13} G when the magnetic field energy density reaches balance with the hydrostatic pressure. Rotation periods vary from $\sim 1 \text{ min}$ to $\leq 100 \text{ yr}$, and a considerable offset between the rotation axis and the magnetic field symmetry axis is often indicated. Simple dipole field structure is evident.

For SS433, we must first obtain an estimate of the energy density in the emission line regions. This is not well determined if the region is truly optically thick because of ambiguity in the number density n . If, however, one assumes that all of the emission eventually escapes in the wings of the lines, a value for n may be obtained [we are indebted to R. L. Brown for pointing this out]. Terlevich and Pringle, apparently following this reasoning, arrived at $n \sim 10^{12} \text{ cm}^{-3}$, although this number is by no means secure except as an upper limit. The presence of large optical depth and the absence of pressure broadening of the lines implies n to be $\sim 10^{10}$ – 10^{12} cm^{-3} , in which case $nkT \sim 10^{-2}$ – 1 in the emission line region. These energy densities may then be compared to the magnetic field strengths at a distance of $r = 10^3 R_* = 10^{12} \text{ cm}$, where R_* is the radius of the magnetic white dwarf. A dipolar surface field of 10^9 G will have a strength of $\sim 1 \text{ G}$ at this radius, and thus the line emission regions at 10^{12} cm can be dynamically controlled by the magnetic field. Gaseous

material has the easiest access onto magnetic field lines in the polar cap regions, and thus we suggest that the emission lines can arise from material forced into near co-rotation at $r \sim 10^{12}$ cm above the two polar caps of a dipole field offset at nearly 90° from the rotation axis.

The velocities required are $\sim 0.15c$, which at $r = 10^{12}$ cm, implies a rotation period of $\sim 1,500$ s, which is well within the observed range for such stars. In order for the star not to be disrupted at this angular velocity, we require $\omega^2 R_* < GM/R_*^2$, or $R_* < 2.5 \times 10^{10} M_\odot^{1/3}$ cm, a condition easily satisfied by a white dwarf. Neglecting for the moment the orbital considerations, material in the polar cap regions will be seen to move towards, or away from the observer with a frequency shift given by

$$v_{\text{em}}/v_{\text{obs}} = \gamma(1 - \beta \cos \theta) \quad (1)$$

where θ is the angle between the velocity vector and the line of sight.

In addition to the co-rotation, the emission line regions also undergo centrifugal acceleration. At $r = 10^{12}$ cm material in the polar cap regions must be accelerated away from, rather than accreted onto the star, since $\omega^2 r = GM/r^2$ at $r = 2.45 \times 10^{10}$ cm. Thus material injected onto field lines at radii less than $10^3 R_*$ will acquire a radical velocity

$$v_r(R) = \int a(r) dt = \int \frac{a(r)}{v_r(r)} dr; \quad a(r) = \omega^2 r - GM/r^2 \quad (2)$$

If injection occurs near the equilibrium radius, then equation (2) yields $v_r \sim \omega r \approx 0.1c$ at $10^3 R_*$, comparable to the co-rotation velocity. Because of this centrifugal acceleration, gas at 10^{12} cm may not be in exact co-rotation if $\rho v_r^2 > B^2/8\pi$. The magnitude of this effect is unknown due to the uncertain values of the gas density and the magnetic field strength at the surface of the white dwarf. The absence of forbidden lines⁹ implies a lower limit to the gas density of about 10^8 cm^{-3} ; a surface field of $\sim 10^{11}$ G will provide exact co-rotation in this case. In either case, velocity shifts of $+50,000 \text{ km s}^{-1}$ and $-30,000 \text{ km s}^{-1}$ are achieved from equation (1) with a value of $\gamma = 1.033$ and a maximum value of $\cos \theta$ of ± 0.5 . Thus the observed velocity extremes result when the plane containing the polar caps is inclined at about 30° to the plane of the sky.

The total amount of gas in the emitting region is small; $m_g \sim 10^{-12} - 10^{-10} M_\odot$, depending on the uncertain value of n , with $10^{-10} M_\odot$ being an upper limit. A similar value for the mass was obtained by Fabian and Rees. Beyond the line emitting region the expanding gas cools in the rapidly diverging field, and the observed line widths result from absorption by the gas of much larger column density which lies above the polar caps; that is, emission is seen only when the poles lie near the sky plane; at other times it is blocked by the larger volume of cooler gas in the line of sight above the polar region facing the observer. The small mass of gas in the line-emitting region can be produced by activity near the surface of the white dwarf triggered by infall of material at lower radii⁸. The infall may arise from direct accretion, or more probably from gas trapped on closed field lines at earlier epochs. In either case ejection is possible only along the open field lines at the poles, and the mass injection rate is $\dot{M} \approx R_{\text{em}}^2 \rho_{\text{em}} v_{r,\text{em}} = 10^{-14} M_\odot \text{ s}^{-1}$, where the subscript 'em' refers to conditions in the emission line region. This small value is at the low end of estimates for accretion onto magnetic white dwarfs¹⁰.

Origin of the periodicity of the emission lines: With the value of $\gamma = 1.033$ needed to obtain the maximum and minimum velocities of the emission lines, the observed periodic nature of these lines follows directly if the plane containing the magnetic poles moves through an angle of $\sim 30^\circ$ with a period of ≈ 160 days. This shifts the plane of rotation into the sky plane with a uniform merging of the emission line velocities to a common value of $+10^4 \text{ km s}^{-1}$ given by the transverse Doppler effect (equation (1), $\theta = 90^\circ$) at $\gamma = 1.033$. For a near 90° offset of magnetic and

spin axes, this implies a 15° precession or nutation angle of the spin vector. During that portion of the period when the pole caps lie in the sky plane, emission is seen from both polar regions. One result of this configuration is that the emission lines will vary in intensity with a period of $\sim 10^3$ s when the lines are at the extremes of their velocity shifts, but that no pulsation will be observed when the lines merge to a common velocity.

Another possibility is that the magnetic field axis is aligned with the spin axis, and that gas is accelerated outwards along the open field lines above the poles. In this case no variation in intensity would be observed, since the line emitting regions would be at all times visible to the observer. Preliminary results from MMT observations¹¹ indicate little variation on short time scales (~ 1 to 100 s) and may favour an aligned spin and magnetic axis model. However, such a model also requires a 15° precession or nutation of the spin vector.

The orbital characteristics of the white dwarf can only be determined indirectly, since its magnitude at 3.5 kpc is $m_v \approx 26 - 29$. A commonly observed configuration is for the spin and orbital angular momenta to be near alignment; this geometry would place the orbital plane of the white dwarf roughly in the sky plane and account for the observed lack of eclipsing of the emission lines. It can easily be shown that precession or nutation driven by the usual gravitational torques upon a non-spherical magnetic dwarf cannot produce the required period, even for extreme differences in the principal moments of inertia¹². To do this would require either an immensely massive primary ($10^6 M_\odot$) or an orbital radius much smaller than the dimension of the line emitting regions. However, either viscous, dissipative coupling or magnetic coupling between the white dwarf and the primary will also exert a torque upon the white dwarf. The maximum co-rotation radius of the magnetic field surrounding the white dwarf is $\sim 10^{13}$ cm, thus an orbital radius of ~ 10 times the size of emission-line regions will provide strong coupling between the stars.

To see if such coupling can produce the required period, we use the standard Euler equations¹³ where the Euler angles ϕ , θ and ψ have their usual meaning. We assume that the star is not grossly deformed, so that the moments of inertia $I_1 \approx I_2 \approx I_3 = I$. If $\vec{T}$ is the total torque exerted by the coupling upon the star, $\vec{T} \approx \vec{F}_d \times \vec{R}$ where $\vec{F}_d$ is the drag force and $\vec{R}$ is the radius of the orbit. A condition of constant precession, namely $\dot{\theta} = \dot{\phi} = 0$, and negligible spin down, $\dot{\psi} \approx 0$, can be found with the required precessional frequency $\dot{\phi} \approx T \sin \theta / (\omega_s I) \sim 10^{-7} \text{ s}^{-1}$, where $\omega_s = \dot{\psi} \sim 10^{-3}$ is the spin of the star and $\theta \sim 15^\circ$. However, in this case the third Euler equation yields $\dot{\phi} \sim 10^{-3} \dot{\phi}$ which implies a very rapidly evolving set of orbital parameters.

A more stable system seems much more probable, and hence we consider nutation ($\dot{\theta} \neq 0$) and slow variation of the other orbital characteristics. For small obliquity, negligible spindown, and a precession period long compared to nutation period, a periodic solution $\dot{\theta} \approx -T\theta/I$ exists, with nutation frequency $=(T/I)^{1/2}$. Setting this equal to the required value of 10^{-7} s^{-1} yields $F_d = 10^{24}$ for $R = 10^{13}$ and a white dwarf mass and radius of $1 M_\odot$ and 10^9 cm, respectively. If the coupling is due to viscous drag, the shear between the co-rotating gas and the atmosphere of the primary will be proportional to the gas density, the square of the relative velocity, and the surface area of the shearing region: $F_d = \epsilon \rho v_{\text{rel}}^2 A$, where $\epsilon \leq 1$ is a measure of the coupling efficiency. For $v_{\text{rel}} \sim 10^{10} \text{ cm s}^{-1}$, $A \sim 10^{24} \text{ cm}^2$, $\epsilon n_{\text{gas}} \approx 10^4$. Thus for gas densities in the outer envelope of the primary as low as 10^6 cm^{-3} , a 1% coupling efficiency will provide the necessary drag force. This efficiency is higher than that normally encountered in purely gas dynamical situations, and its enhancement may be attributed to magnetic coupling at $R = 10^{13}$ cm.

Another contribution to coupling between the stars arises from interaction between the magnetic fields of the primary and secondary. The luminosity and emission line spectra observed at $+180 \text{ km s}^{-1}$ imply that the primary is a fairly early type star with mass $\sim 1 - 10 M_\odot$. If, for example, the primary were an Ap star with surface field $\sim 30,000$ G and radius 5×10^{11} cm, the

magnetic torque induced upon the white dwarf is $T \sim (4\pi/3)B_2 r_2^3 B_1 (r_1/R)^3 \sim 1 \times 10^{37}$, where the subscripts 1 and 2 refer to the primary and to the white dwarf, respectively. The resulting drag force is $F_d = T/R \sim 10^{24}$ dyn, which is the required value. If such coupling is dominant it would imply that SS433 is similar to the model of AM Her suggested by Stockman *et al.*¹⁴. Radio emission can easily arise in the lower regions of the white dwarf magnetosphere, with the X-ray emission being produced by infall onto the stellar surface⁸.

We thank R. L. Brown, J. Liebert, R. O'Connell, S. O'Dell and R. Thompson for helpful discussion, and B. Margon for data. The National Radio Astronomy Observatory is operated by the Associated Universities Inc. under contract with the NSF and Kitt Peak National Observatory by the Association of

Universities for Research in Astronomy Inc. under NSF contract.

Received 21 May; accepted 30 July 1979.

1. Margon, B. *et al. Astrophys. J. Lett.* **230**, L41 (1979).
2. Margon, B. *et al. IAU Circ. No.* 3345 (1979).
3. Liebert, J. *et al. Astrophys. J.* (in the press).
4. Abell, G. O. & Margon, B. *Nature* **279**, 701–703 (1979).
5. Seaquist, E. R. *et al. Preprint* (1979).
6. Terlevich, R. J. & Pringle, J. E. *Nature* **278**, 719 (1979).
7. Fabian, A. C. & Rees, M. J. *Mon. Not. R. astr. Soc.* **187**, 13P (1979).
8. Angel, J. R. A. *Rev. Astr. Astrophys.* **16**, 487 (1978).
9. Liebert, J. *et al. Nature* **279**, 384–387 (1979).
10. Fabian, A. C. *et al. Mon. Not. R. astr. Soc.* **173**, 43 (1976).
11. McGraw, J. *et al. IAU Circ. No.* 3363 (1979).
12. Routh, E. J. *Dynamics of a System of Rigid Bodies*, 340–365 (Dover, New York, 1955).
13. Goldstein, H. *Classical Mechanics*, 182 (Addison, New York, 1950).
14. Stockman, H. S. *et al. Astrophys. J.* **217**, 815 (1977).

Oxidation of CS₂ and COS: sources for atmospheric SO₂

Jennifer A. Logan, Michael B. McElroy, Steven C. Wofsy & Michael J. Prather

Center for Earth and Planetary Physics, Harvard University, Cambridge, Massachusetts 02138

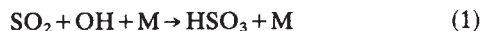
Oxidation of CS₂ and COS initiated by reaction with OH can provide a source of atmospheric SO₂ as large as 12 Mtonnes S yr⁻¹ and may represent the dominant source of SO₂ in remote regions of the marine troposphere.

EMISSION of sulphur associated with combustion of fossil fuel represents a major source of atmospheric pollution^{1,2}, contributing ~65 Mtonnes S yr⁻¹. Although the chemistry of sulphur in polluted atmospheres has been studied extensively, our knowledge of the natural cycle remains rudimentary. Measurements^{3,4} at remote marine locations in the Southern Hemisphere indicate that the atmosphere contains a relatively constant background level of COS and SO₂. We shall argue here that SO₂ is derived at least in part from oxidation of COS, with additional sources due to CS₂ and dimethyl sulphide (DMS).

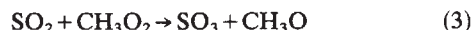
Sulphurous compounds in the atmosphere

It seems probable that the chemistry of sulphurous compounds in the troposphere should be dominated by reactions involving OH⁵⁻⁷. The hydroxyl radical is formed by reaction of O(¹D) with H₂O, and is removed primarily by reaction with CO and CH₄. Results of model calculations for the concentration of OH are given in Fig. 1. We believe that concentrations shown here should be reliable to within about a factor of 2, and uncertainties of this magnitude should not qualitatively affect the arguments which follow.

Sulphur dioxide may be removed from the atmosphere by various heterogeneous processes⁸, and by gas phase reactions^{5,9} involving OH, HO₂ and CH₃O₂:



and



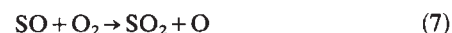
The products in reactions (2) and (3) have not been measured directly, and these reactions may possibly involve formation of adducts. Reaction (1) has been studied by several groups^{5,10}, using a variety of techniques, and its rate constant is probably known to within about a factor of 2. It constitutes a major, perhaps dominant, path for removal of SO₂. Reaction (3) could be comparable to reaction (1). Considerable uncertainty attaches, however, not only to the mechanism of reaction (3), but also to the concentration of atmospheric CH₃O₂. Reaction (2) is of minor importance.

Studies of SO₂ in polluted air parcels^{11,12} suggest an atmospheric lifetime of about a day. Calculations presented here for the marine atmosphere, using the kinetic data summarised in Table 1 but omitting heterogeneous reactions, indicate a lifetime of about 8 days. The difference may reflect in part the role of heterogeneous chemistry, in part omission of processes in polluted air which might be expected to result in higher concentrations of radicals (RO, RO₂). The calculations presented here were carried out using a one-dimensional photochemical model, as described by Logan *et al.*¹³, with modifications to their chemical scheme as given in Table 1. Vertical profiles for O₃, CO, CH₄ and H₂O were specified using available observational data and calculations were performed for latitudes 45°N, 15°N, 15°S and 45°S.

Significant quantities of SO₂ may be formed during oxidation of COS and CS₂. It appears⁶ that oxidation of COS should be initiated by reaction with OH,



Subsequent steps may involve production and removal of SO:



Rate constants for reactions (5)–(7) have not been determined experimentally. However, it seems reasonable that SH should be oxidised in the atmosphere to SO₂ by reactions (5)–(7) or by some other reaction mechanism. Measurements by Kurylo of the rate constant for reaction (4) and for the corresponding reaction with CS₂ are in conflict with results reported by Atkinson *et al.*¹⁴. Sources of the discrepancies are discussed in detail in ref. 6, and are attributed to complications due to photolysis of COS and to the occurrence of secondary reactions. Kurylo's experiments were performed in conditions designed to minimise these problems and his rate constants have been adopted accordingly. We estimate a global mean lifetime for COS of 200 days. Combining this lifetime with measured concentrations of COS^{4,15}, 0.5 ± 0.1 p.p.b. (parts per 10⁹, by volume) we infer a global destruction rate for COS equivalent to 4 ± 0.8 Mtonnes S yr⁻¹.

Oxidation of CS₂ can provide a significant source of COS^{6,7}, through



Subsequent oxidation of SH, through reactions (5) and (6),

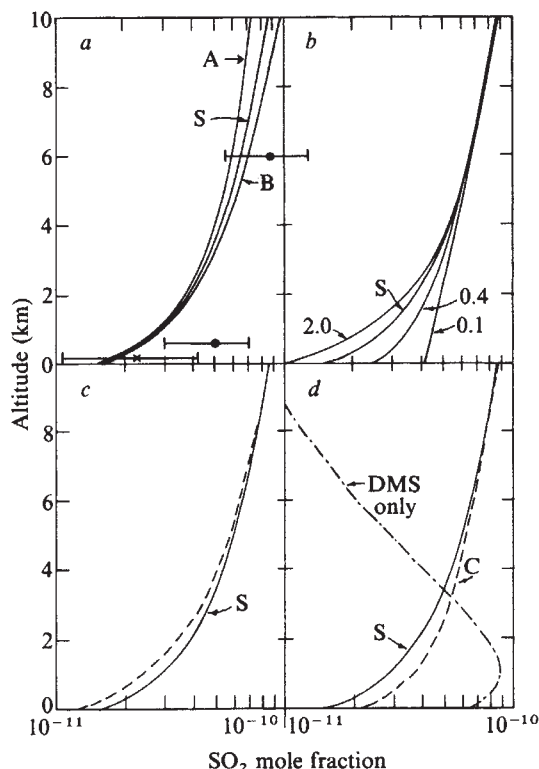


Fig. 1 Altitude profiles for OH and O₃ at 15°S latitude at equinox. *a*, Diurnally averaged concentrations for OH. The model calculations follow the time-dependent variation of short lived chemical species resulting from the diurnally varying radiation field which includes Rayleigh scattering. Calculations are shown for clear sky conditions (A) and a combination of clear sky (70%) and fully reflecting cloud layers at 3 km (17%) and 5 km (13%) (B). Vertical transport is simulated by 'eddy diffusion', with a diffusion coefficient equal to $10^5 \text{ cm}^2 \text{ s}^{-1}$. The chemical model of Logan *et al.*¹³ has been updated as described in Table 1. Fixed profiles for O₃⁴⁸ (see *b*), CO⁴⁹ (0.6 p.p.b.), CH₄⁵⁰ (1.5 p.p.b.) and H₂O (taken from the relative humidity values of Rasmusson⁵¹) were used. The mixing ratio of NO + NO₂ was taken to be 10^{-11} at 0 km, (ref. 52) and a deposition velocity of 1 cm s^{-1} was used for HNO₃. Note that the model implies a value of $1 \times 10^6 \text{ molecules cm}^{-3}$ for the globally averaged concentration of tropospheric OH.

should provide an additional source for SO₂. We estimate a lifetime for CS₂ due to reaction (8) of 60 days. Observational data for CS₂ are sparse, confined to the UK¹⁵, and indicate a range of mixing ratios between 0.07 and 0.37 p.p.b. Sandalls and Penkett¹⁵ point out that their data are consistent with the view that the ocean and atmosphere are in equilibrium with respect to CS₂. Using marine measurements of CS₂ by Lovelock¹⁶, and solubility data by Seidell¹⁷, they inferred an equilibrium concentration for the atmosphere equivalent to 0.22 ± 0.06 p.p.b. We adopt this value for the mixing ratio of CS₂ to represent global mean conditions. Used in conjunction with the chemical model of Table 1 and with the OH data of Fig. 1, it implies a global destruction rate for CS₂ equal to $12 \pm 3 \text{ Mtonnes S yr}^{-1}$. The associated sources for COS and SO₂ are $6 \pm 1.5 \text{ Mtonnes S yr}^{-1}$ and $12 \pm 3 \text{ Mtonnes S yr}^{-1}$ respectively. Note that the source for COS due to oxidation of CS₂ could account for the bulk of atmospheric COS as discussed by Sze and Ko⁷.

Source rates quoted here are directly proportional to values calculated for the concentration of OH. Studies of CH₃CCl₃^{18,19} suggest that concentrations of OH given in Fig. 1 may be too high by about a factor of 2. If valid, this analysis would require corresponding reduction in source strengths inferred for CS₂ and COS. Relative abundances of CS₂, COS and SO₂ in remote locations are comparatively insensitive to details of the chemical model for OH, and provide a useful test of the hypothesis

proposed here, that is, that oxidation of CS₂ and COS might represent dominant background sources of SO₂. Our model predicts concentration ratios CS₂:COS:SO₂ at 6 km of 0.3:1.0:0.13 which may be compared with observed ratios⁴ of 0.44:1.0:0.17. This is a fairly good agreement, when one remembers that the observational data for COS and SO₂ are uncertain to within about $\pm 30\%$, that our value for the concentration of CS₂ was inferred from measurements of dissolved CS₂, and that the temperature dependencies of the rate constants for the key reactions (1), (4) and (8) are not well defined. In our model, one-half of the SO₂ source is derived from oxidation of CS₂, for which observational data are extremely limited. The mixing ratio for COS is fairly well defined and, given the uncertainties in the rate constants for reactions (1), (3) and (4), oxidation of COS alone might provide the major source for background SO₂.

Results and discussion

Calculated height profiles for SO₂ are shown in Fig. 2, which includes comparison with the observations by Maroulis *et al.*⁴ and Nguyen Ba Cuong *et al.*²⁰. We define a standard model, S, characterised by rate constants and parameters summarised in Table 1. Our choice of rate constant for reaction (1) is based on data reported by Castleman *et al.*²¹ and Cox²². In their studies the rate for reaction (1) was measured relative to the rate for reaction of OH with CO. We applied suitable correction to account for the known pressure dependence²³ of OH + CO.

Our standard model assumes that the rate for reaction (1) is independent of temperature. Dependence of results on choice of rate expressions for reaction (1) is shown in Fig. 2*a*, which also includes models allowing modest negative activation energies

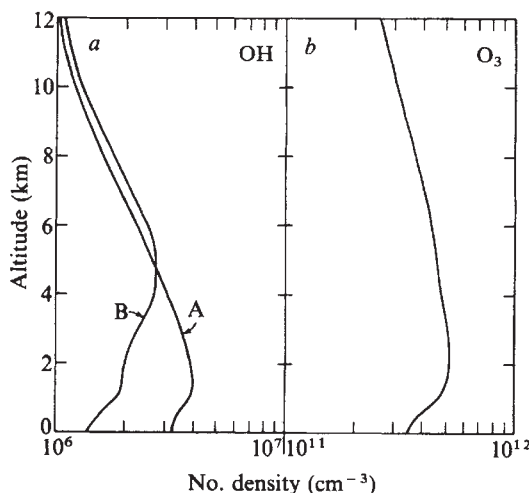


Fig. 2 Altitude profiles for SO₂ at 15°S latitude at equinox. The profiles were calculated with the model given in Fig. 1, curve B. The curves labelled S used the standard chemical model of Table 1, with a deposition velocity for SO₂ of 1 cm s^{-1} and a zero-flux boundary condition at the tropopause. The COS mixing ratio was specified at 0.5 p.p.b. Data are from Maroulis *et al.*⁴ (●), aircraft flight over the Pacific Ocean, 0–67° S; and Nguyen Ba Cuong *et al.*²⁰ (x), Indian Ocean cruise, 5–25° S. *a*, The effect of temperature dependent rate constants: curve A, $k_b (\text{OH} + \text{SO}_2) = 5.7 \times 10^{-33} \exp(1,000/T)$; curve B, $k_4 = 3.4 \times 10^{-14} \exp(150/T)$, $k_8 = 1.1 \times 10^{-13} \exp(150/T)$. *b*, Uses the standard model S, but with values for the deposition velocity for SO₂ equal to 2.0, 0.4 and 0.1 cm s^{-1} as noted. *c*, The dashed curve includes loss of SO₂ to aerosols, with a sticking efficiency of 10^{-3} , as described in the text. *d*, The effect of a source of SO₂ from oxidation of DMS. Curve C includes an ocean source of DMS of $2.5 \text{ Mtonnes S yr}^{-1}$, in addition to SO₂ sources from CS₂ and COS oxidation, and may be considered as our best model. The curve labelled 'DMS only' omits all sources of SO₂ except that from DMS oxidation, and allows for a DMS source of $25 \text{ Mtonnes S yr}^{-1}$, the value suggested recently by Nguyen Ba Cuong *et al.*³⁷. The calculated profiles of SO₂ exhibit small (<15%) variations with latitude between 45° N and 45° S.

for reactions (4) and (8). It seems probable that rates for reactions (4) and (8) should be faster at low temperature, given the reaction mechanism proposed by Kurylo^{6,24}. The temperature dependence of k_1 is poorly determined. The rate expression given in Fig. 2a was based on preliminary work by Castleman and Tang²⁵ and Erler *et al.* (ref. 26 and results in ref. 10).

Figure 2b shows the dependence of calculated profiles for SO₂ on the choice of deposition velocity, while Fig. 2c shows the possible influence of heterogeneous loss due to aerosols. The general behaviour of the profiles in Fig. 2a–c is similar. In all cases we find higher mixing ratios at high altitude, consistent with the observational data of Georgii *et al.*²⁷ and Maroulis *et al.*⁴. Uncertainties in loss processes tend to affect mainly concentrations computed for the lowest altitudes. Deposition velocities over land⁸ are reported in the range 0.1–2.0 cm s^{−1} while Hicks and Liss²⁸ suggest that a value of about 1 cm s^{−1} should be appropriate for moderate wind conditions at sea.

Results in Fig. 2c imply that loss of SO₂ to aerosols should be relatively unimportant, at least in the marine atmosphere. We used an integrated surface area for aerosols equivalent to 1×10^{-7} cm² per cm³, an estimate based on particle size distributions measured by Meszaros and Vissy²⁹, Blifford³⁰, and Jaenicke *et al.*³¹. We investigated sticking efficiencies for SO₂ as high as 10^{-3} , although experimental data³² suggest efficiencies much lower than this, in the range 10^{-3} – 10^{-6} . Results in Fig. 2 do not allow for transient reductions in SO₂ which might arise due to rainfall.

We assumed a rate constant for reaction (3) equal to 10^{-15} cm³ s^{−1}. With this choice, reaction (3) represents a relatively minor sink for SO₂. Measurements by Kan *et al.*³³ indicate removal of CH₃O₂ by SO₂, with an effective rate constant of 5 – 10×10^{-15} cm³ s^{−1}. If loss of CH₃O₂ observed in the laboratory is associated with the reaction channel (3), concentrations in Fig. 2 could be reduced by about 20%. This estimate assumes, however, an unproven model for oxidation of atmospheric CH₄. It seems unlikely that k_3 should be as large as 10^{-14} cm³ s^{−1} given the much smaller value, $\leq 10^{-17}$ cm³ s^{−1} (refs 34, 58), reported for the analogous reaction (2).

Oxidation of DMS could provide an additional source of SO₂ as suggested by Lovelock³⁵ and discussed further by Maroulis and Bandy³⁶ and Nguyen Ba Cuong *et al.*³⁷. The associated

source of SO₂ would be confined to the lowest 2 km of the atmosphere as the reaction of DMS with OH is exceedingly rapid³⁸. Thus DMS cannot account for the height profile observed for SO₂, a point shown in Fig. 2d. Similar arguments apply to the suggestion that oxidation of H₂S might produce a large source of SO₂. Hydrogen sulphide also reacts very rapidly with OH³⁹, and concentrations as low as a few parts per trillion have been observed in tropical marine air⁴⁰. Most estimates of the contribution of H₂S to the atmospheric sulphur cycle have been derived by subtracting other terms in the budget for sulphur, and are consequently highly uncertain^{1,2}. While biogenic sources of H₂S and DMS may provide significant contributions to the atmospheric sulphur cycle in selected environments (for example, near swamps or salt marshes), these sources cannot account for the observations of SO₂ discussed here.

Liss and Slater⁴¹, using measurements by Lovelock *et al.*³⁵, estimate a net source of 3.7 Mtonnes S yr^{−1} associated with release of marine DMS, a source whose magnitude should be reduced by about a factor of 1.5 to account for more recent estimates of rates for gas exchange across the air–sea interface⁴².

It seems probable that the concentrations of atmospheric CS₂ and COS should reflect a combination of biogenic and industrial sources^{43,44}. Preliminary measurements of the emission rates of sulphurous gases over salt marshes have shown that fluxes of COS might be comparable to, and in some cases greater than fluxes of H₂S⁴⁵. The relatively uniform distribution^{4,13,46,47} observed for COS suggests a globally distributed source. This apparently argues against a significant industrial component, although there might be important production due to forest fires and combustion associated with land clearance for agriculture. The possibility of a dominant industrial component for CS₂ cannot be excluded, given limitations in the present data base. It seems more likely, however, that diffuse sources should dominate for CS₂ as well as for COS. These sources may include a component due to combustion although thermodynamic considerations suggest that reactions in reducing environments may be more important. It seems probable that these reactions should be microbially mediated: swamps, anaerobic muds and the digestive tracts of various animals offer plausible sites. The relatively uniform distribution observed for SO₂ in the upper marine troposphere⁴ requires a diffuse atmospheric source for

Table 1 Kinetic data

No.	Reaction	Rate expression*	Refs and notes
1	OH + SO ₂ + M → HSO ₃ + M	$k_a = 6.8 \times 10^{-13}$ $k_b = 1.6 \times 10^{-31}$ $1/k = 1/k_a + 1/k_b M$ $k_1 = k \times (k_9/1.4 \times 10^{-13})$	Fit to the data of refs 21, 22
2	HO ₂ + SO ₂ → SO ₃ + OH	2.0×10^{-17}	34
3	CH ₃ OO + SO ₂ → products	1.0×10^{-15}	See text
4	OH + COS → SH + CO ₂	5.66×10^{-14}	6
7	OH + CS ₂ → SH + COS	1.85×10^{-13}	6
8	OH + (CH ₃) ₂ S → products	$6.08 \times 10^{-12} \exp(134/T)$	38
9	OH + CO → CO ₂ + H	$2.10 \times 10^{-13} f(M) \exp(-115/T)$ $f(M) = (1 + 3.7 \times 10^{-20} \times M)$	53 23
10	HO ₂ + O ₃ → OH + 2O ₂	$1.40 \times 10^{-14} \exp(-580/T)$	Zahniser and Howard (personal communication)
11	HO ₂ + NO → OH + NO ₂	$3.3 \times 10^{-12} \exp(254/T)$	54
12	HO ₂ + HO ₂ → H ₂ O ₂ + O ₂	$3.9 \times 10^{-14} f(\text{H}_2\text{O}) \exp(1245/T)$ $f(\text{H}_2\text{O}) = 1 + 2.5 \times 10^{-18} \times [\text{H}_2\text{O}]$	55, 56
13	H ₂ O ₂ + $h\nu$ → OH + OH		57
14	CH ₃ OO + NO → CH ₃ O + NO ₂	8.0×10^{-12}	Estimated
15	CH ₃ OO + HO ₂ → CH ₃ OOH + O ₂	3.0×10^{-12}	Estimated
16	H ₂ O ₂ heterogeneous loss (units, s ^{−1})	1.16×10^{-5} 0–4 km $1.16 \times 10^{-5} \exp(1.6 - 0.4 z)$ 4–12 km	
17	SO ₂ + aerosol → heterogeneous loss (probability)	0.0	†

* Units cm³ s^{−1} for bimolecular reactions; cm⁶ s^{−1} for termolecular reactions. The complete chemical scheme is described in Logan *et al.*¹³.

† The integrated surface area for marine aerosols was taken to be 1.0×10^{-7} cm² per cm³, and a scale height for the aerosol distribution function of 1 km was adopted (see refs 29–31).

this gas. Oxidation of COS and perhaps CS₂ as discussed here can provide a relatively straightforward explanation for the existence of a uniform background level of SO₂.

Received 17 May; accepted 20 July 1979.

1. Friend, J. P. in *Chemistry of the Lower Atmosphere* (ed. Rasool, S. I.) (Plenum, New York, 1973).
2. Granat, L., Rodhe, H. & Hallberg, R. O. *Ecol. Bull.* **22**, 89–134 (1976).
3. Mészáros, E. *Atmos. Envir.* **12**, 699–705 (1978).
4. Maroulis, P. J., Torres, A. L., Goldberg, A. B. & Bandy, A. R. *Trans. Am. geophys. Un* **59**, 1081–1082 (1978); 1082 (1978).
5. Calvert, J. G., Su, F., Bottenheim, J. W. & Strausz, O. P. *Atmos. Envir.* **12**, 197–226 (1978).
6. Kurylo, M. J. *Chem. phys. Lett.* **58**, 238–242 (1978).
7. Sze, N. D. & Ko, M. K. W. *Nature* **280**, 308–310 (1979).
8. Garland, J. A. *Atmos. Envir.* **12**, 349–362 (1978).
9. Cox, R. A. *Phil. Trans. R. Soc. A* **290**, 543–550 (1979).
10. Moortgat, G. K. & Junge, C. E. *Pure appl. Geophys.* **115**, 759–774 (1977).
11. Prahm, L. P., Torp, U. & Stern, R. M. *Tellus* **28**, 355–372 (1976).
12. Nguyen, B. C., Bonsang, B., Lambert, G. & Pasquier, J. L. *Pure Appl. Geophys.* **113**, 489–500 (1975).
13. Logan, J. A., Prather, M. J., Wofsy, S. C. & McElroy, M. B. *Phil. Trans. R. Soc. A* **290**, 187–234, 1978.
14. Atkinson, R., Perry, R. A. & Pitts, J. N., Jr *Chem. phys. Lett.* **54**, 14–18 (1978).
15. Sandalls, F. J. & Penkett, S. A. *Atmos. Envir.* **11**, 197–199 (1977).
16. Lovelock, J. E. *Nature* **248**, 625–626 (1974).
17. Seidell, A. in *Solubilities of Inorganic and Metal Compounds, Vol. 1* (Van Nostrand, New York, 1940).
18. Singh, H. B. *Geophys. Res. Lett.* **4**, 453–456 (1977).
19. Lovelock, J. E. *Nature* **267**, 32 (1977).
20. Nguyen, B. C., Bonsang, B. & Lambert, G. *Tellus* **26**, 241–248 (1974); Cuong, B. C., Bonsang, B., Pasquier, J. L. & Lambert, G. *J. Res. Atmos.* **8**, 831–844 (1974).
21. Castleman, A. W., Davis, R. E., Munkelwitz, H. R., Tang, I. N. & Wood, W. P. *Int. J. Chem. Kin. Symp.* **1**, 629–640 (1975).
22. Cox, R. A. *Int. J. Chem. Kin. Symp.* **1**, 379–398 (1975).
23. Biermann, H. W., Zetzsch, C. & Stuhl, F. *Ber. Bunsenges. Phys. Chem.* **82**, 633–639 (1978).
24. Kurylo, M. J. & Laufer, A. H. *J. chem. Phys.* **70**, 2032–2033 (1979).
25. Castleman, A. W. & Tang, I. N. *J. Photochem.* **6**, 349–354 (1976).
26. Erler, K., Field, D. & Zellner, R. *4th Int. Conf. on Gas Kinetics Edinburgh* (1975).
27. Georgii, H. W. & Gravenhorst, G. *Pure appl. Geophys.* **115**, 503–511 (1977).
28. Hicks, B. B. & Liss, P. S. *Tellus* **28**, 349–354 (1976).
29. Mészáros, A. & Vissy, K. *J. atmos. Sci.* **5**, 101–110 (1974).

We thank N. Dak Sze for helpful discussions, and acknowledge support from NASA, grant NSG-2031, and the NSF, grant ATM 78-09712.

30. Blifford, I. H. *J. geophys. Res.* **75**, 3099–3103 (1970).
31. Jaenicke, R., Junge, C. & Kanter, H. J. *Met. Forsch. Ergebn.* **87**, 1–54 (1971).
32. Judeikis, H. S., Stewart, T. B. & Wren, A. G. *Atmos. Envir.* **12**, 1633–1641 (1978).
33. Kan, C. S., McQuigg, R. D., Whitbeck, M. R. & Calvert, J. G. *Int. J. Chem. Kin.* (in the press).
34. Burrows, J. P., Cliff, D. I., Harris, G. W., Thrush, B. A. & Wilkinson, J. P. T. *W.M.O. Tech. Not. No.* **511**, 25–28 (1978).
35. Lovelock, J. E., Maggs, R. J. & Rasmussen, R. A. *Nature* **237**, 452–453 (1972).
36. Maroulis, P. J. & Bandy, A. R. *Science* **196**, 647–648 (1977).
37. Nguyen, B. C., Gaudry, A., Bonsang, B. & Lambert, G. *Nature* **275**, 637–639 (1978).
38. Kurylo, M. J. *Chem. phys. Lett.* **58**, 233–237 (1978).
39. Perry, R. A., Atkinson, R. & Pitts, J. N. *J. chem. Phys.* **64**, 3237–3239 (1976).
40. Slatt, B. J., Natusch, D. F. S., Prospero, J. M. & Savoie, D. L. *Atmos. Envir.* **12**, 981–991 (1978).
41. Liss, P. S. & Slater, P. G. *Nature* **247**, 181–184 (1974).
42. Broecker, W. S. & Peng, T. -H. *Tellus* **26**, 21–35 (1974).
43. Crutzen, P. J. *Geophys. Res. Lett.* **3**, 73–76 (1976).
44. Peyton, T. O., Steele, R. V. & Mabey, W. R. *Stanford Res. Inst. Rep. No.* **68-01-2940** (1976).
45. Aneja, V. P., Overton, J. H., Cupitt, L. T., Durham, J. L. & Wilson, W. E. *Tellus* **31**, 174–178 (1979).
46. Hanst, P. L., Spiller, L. L., Watts, D. M., Spence, J. W. & Miller, M. F. *J. Air Pollut. Contr. Ass.* **25**, 1220–1226 (1975).
47. Maroulis, P. J., Torres, A. L. & Bandy, A. R. *Geophys. Res. Lett.* **4**, 510–512 (1977).
48. Chatfield, R. & Harrison, H. J. *geophys. Res.* **82**, 5969–5976 (1977).
49. Heidt, L. E., Kransnec, J. P. *Trans. Am. geophys. Un* **59**, 1077 (1978).
50. Ehhalt, D. H., Schmidt, U. *Pure appl. Geophys.* **116**, 452–464 (1978).
51. Rasmusson, E. M. in *The General Circulation of the Tropical Atmosphere and Interactions in the Extraterrestrial Latitudes* (ed. Newall, R. E., Ch. 5. MIT Press, 1972).
52. McFarland, M., Kley, D., Drummond, J. W., Schmeltekopf, A. L. & Winkler, R. H. *Geophys. Res. Lett.* **6**, 605–608 (1979).
53. Greiner, N. R. *J. chem. Phys.* **51**, 5049–5051 (1969).
54. Howard, C. J. & Zahniser, M. S. *J. chem. Phys.* (submitted).
55. Cox, R. A. *W.M.O. Tech. Note No.* **511**, 17–27 (1978).
56. Hamilton, E. J. & Lii, R.-R. *Int. J. chem. Kin.* **9**, 875–885 (1977).
57. Molina, L. T., Schinke, S. D. & Molina, M. J. *Geophys. Res. Lett.* **4**, 580–582 (1977).
58. Graham, R. A., Winer, A. M., Atkinson, R. & Pitts, J. N. *J. Phys. Chem.* **83**, 1563–1567 (1979).

The effects of past variations of the Earth's rotation rate on climate

B. G. Hunt

Australian Numerical Meteorology Research Centre, PO Box 5089AA, Melbourne, Australia 3001

Palaeological evidence indicates that the rotation rate of the Earth during the late Precambrian was 2–2.5 times faster than now. Typically high rotation rate in a fluid system reduces the characteristic size of dynamic features, their associated transport processes and the overall intensity of the motions. The climatic consequences of the high rotation rate during the Precambrian are outlined. In particular, it is hypothesised that variations in the rotation rate imposed an important, and possibly dominant, effect on the genesis and termination of the Precambrian ice age. Other possible implications of the high rotation rate to oceanography, life forms and geology are briefly noted.

POTENTIAL mechanisms for producing climatic variations include solar perturbations owing to orbital variations of the Earth¹, volcanic activity², tectonic movements³, fluctuations in solar output⁴, glacier surgings⁵ and oceanic feedback⁶. On a time scale of hundreds of millions of years a further factor, which has received relatively little attention⁷, is the increase in the rotational period of the Earth. This increase is very small (~2.5 ms per century) and is still taking place. However, since the atmospheric and oceanic circulations are known to be critically influenced by the Earth's rotation rate it follows that climate should also be affected. Hence it is of value to consider what

impact past variations in the rotation rate might have had on the evolution of the Earth's palaeoclimate, and to see how well the deductions fit the known facts.

This assessment will be obtained by interpreting the results of a numerical model of the atmospheric circulation. This general circulation model of the atmosphere was hemispheric in extent, devoid of topography and thus oceanic influences, and was set up for fixed annual mean conditions⁸. Despite these apparently serious omissions the model is capable of a remarkable simulation of many features of the atmosphere^{8,9}, which encourages the belief that its application to problems concerned with variations in the rotation rate of the Earth has merit. This model has now also been run with the rotation rate increased and decreased by a factor of 5 as part of a separate study¹⁰, and it is this data base of three experiments which will be used here.

Prehistorical perspective

The rotational history of the Earth can be traced back to the late Precambrian era, about 1.5×10^9 yr ago. Based on growth patterns from shells and corals the Earth then had at least 800–900 solar days per yr, implying a rotation rate 2–2.5 times greater than the present value¹¹. MacDonald¹² has calculated the past variation of the rotation rate and his results for one set of assumptions are shown in Fig. 1. This illustrates a possible increase in rotation rate with decreasing time up to about

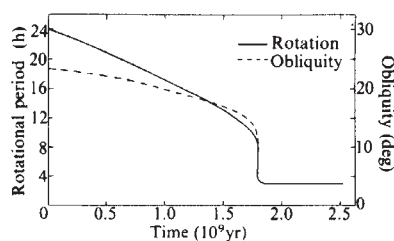


Fig. 1 Variation with time of Earth's rotational period (left-hand scale) and obliquity (right-hand scale).

1.7×10^9 yr ago, when a discontinuity associated with the closest approach of the Moon to the Earth may have occurred. Such a discontinuity, if real, would have created catastrophic climatic and other changes. Consequently the emphasis here will be placed on the less dramatic, but continuous, climatic effects associated with the subsequent steady decrease of the rotation rate between 1.7×10^9 yr ago and the present.

It is also necessary to consider other factors which may have been important to climatic evolution in this time frame in order to clarify the contribution from rotation rate variations. Thus continental drift, as recently reconstructed by Irving¹³ back to the middle Devonian (375×10^6 yr ago), could markedly affect climate. This can occur via continentality effects associated with large land masses¹⁴, or through changes to ocean depths or basin sizes¹⁵. In particular shallower oceans would have had a less ameliorating influence on climatic developments. (Note that rotation rate variations would also have affected the dynamics of the motions in the molten interior of the Earth, and therefore geomagnetic activity, in addition to the motions of the atmosphere and oceans. Presumably such dynamic variations could have influenced the large-scale convective motions responsible for tectonic activity and volcanism, and this could be an important factor in the geological evolution of the Earth.) A further prehistorical factor is the change in the obliquity of the Earth's Equator to the ecliptic associated with the dynamics of the Earth-Moon system (Fig. 1). The lower obliquity in the past implies less annual mean solar input to high latitudes and therefore possibly colder polar regions, all else being the same. The luminosity of the Sun has also probably increased with time; a 25% increase over the past 4.5×10^9 years has been suggested¹⁶. Wetherald and Manabe¹⁷ have shown that 1–2% variations in the solar constant could have significant effects on the climate of a general circulation model, particularly at high latitudes. Composition changes in the atmosphere during its evolution are also relevant, and contribute to a computed 10 K increase in mean surface temperature over the past 1.5×10^9 yr (ref. 16). The principal change in composition during this time span is supposedly in the O_2/N_2 ratio. A further important factor in climatic variability is volcanic activity². Other forcing functions could be enumerated but not all of these are relevant to the time frame of interest here. The above list is thought to contain the principal precursors of climatic change which should be considered simultaneously with an evaluation of the contribution associated with the decrease in the Earth's rotation rate.

Model climatology at high rotation rate

A summary of the results of the general circulation model run with a fivefold increase in rotation rate¹⁰ is given here to illustrate the markedly different atmospheric behaviour in these conditions. For simplicity this model was run with all radiative terms (including cloud cover, surface albedo and CO_2 amount) set at current values. These and other model parameterisations are based on present atmospheric conditions which would not necessarily apply in the extreme conditions considered here. This undoubtedly distorted the details of the results but the dynamically induced changes in the large scale circulation are reasonably realistic.

An important difference is shown in Fig. 2, where a comparison of near surface (0.9 km) synoptic meridional velocity distributions for this model and the control (normal rotation rate) reveals a distinct reduction in scale size at high rotation rate. Wavenumber 13 predominates compared to wavenumber 6 for the control. This result is not resolution limited¹⁸, and in fact the overall behaviour of the model at all three rotation rates closely resembled the general flow characteristics produced by laboratory analogues of the atmosphere in appropriate conditions^{19,20}. A consequence of the smaller scale size was that the synoptic distributions of the various atmospheric fields (velocity, temperature, moisture) were poorly correlated. This resulted in a reduced tropospheric polewards transport of sensible and

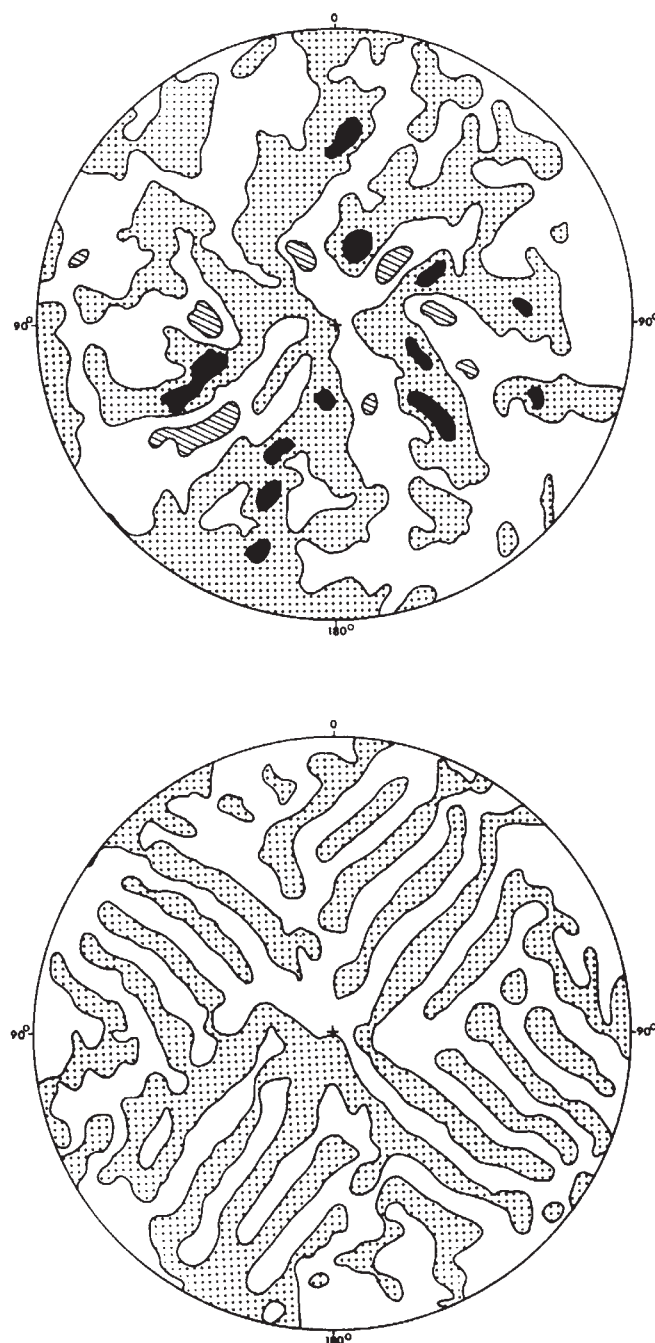


Fig. 2 Instantaneous synoptic meridional (north-south) velocity distributions for normal rotation and fivefold rotation rates are given in top and bottom of figure respectively. The circumscribing circle is the Equator, the + sign the Pole. Solid areas, $> 10 \text{ m s}^{-1}$; stippled, $0-10 \text{ m s}^{-1}$; blank, $0-(-10) \text{ m s}^{-1}$; cross-hatched, $> -10 \text{ m s}^{-1}$.

latent heat by the dominant large-scale eddies, thus causing a much colder high latitude region and a warmer tropics.

The mean meridional stream function increased from three to five cells, with the tropical Hadley cell being markedly reduced in latitudinal extent and intensity. The mean zonal winds were almost halved in intensity, with the tropospheric westerly jet stream being located at 10° latitude. The stratospheric westerly jet virtually ceased to exist as a separate entity. A simple application of the thermal wind equation suggests that a more intense tropospheric jet should have been obtained because of the increased tropospheric latitudinal temperature gradient in the model. However, this equation is invalid at fast rotation rates where motions are highly ageostrophic, and again the model's behaviour is confirmed by laboratory analogue experiments^{19,20}.

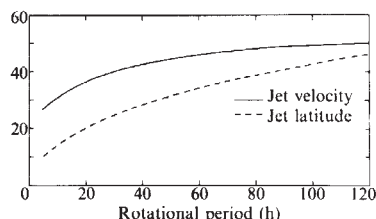


Fig. 3 Variation with Earth's rotational period of the intensity of the mean zonal wind (m s^{-1}) of the tropospheric jet, and of the location of the jet (deg latitude).

The conventionally defined arid region of the globe narrowed in latitudinal extent and moved towards the Equator, but most of the globe polewards of 45° also became semi-arid as the increased coldness of the troposphere caused a general decline in evaporation and precipitation in this region.

Implications of rotation rate variations for Earth's climate

Consider first the overall impact of the change in wind strength with rotation rate as derived from the three model experiments. This is illustrated in Fig. 3 by the variation of the intensity and latitude of the tropospheric jet stream, which is located at a height of about 10 km. This indicates that for a rotational period of 9 h (late Precambrian era) the jet stream intensity was reduced by about 20% compared to its current value, but, more importantly, the jet was located much closer to the Equator at 15° latitude. As illustrated in Fig. 1 of ref. 10 the confinement of the jet to low latitudes was associated with a reduction in the overall intensity of the zonal winds in the troposphere as a whole, to a much greater extent than that implied by the 20% reduction in jet intensity. A prime consequence of this constraint is that in Precambrian times the surface wind stress, which is normally proportional to the square of the wind velocity, would have been reduced by possibly as much as a factor of 2. A rough indication of this change can be obtained by comparing the intensity and variability of the near surface synoptic wind distributions in Fig. 2.

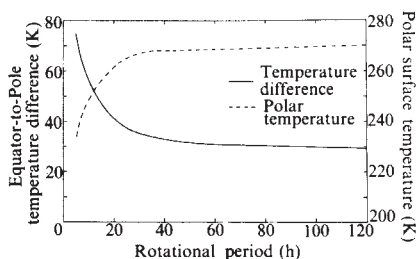


Fig. 4 Variation with Earth's rotational period of the Equator-to-Pole temperature difference and polar surface temperature.

Perhaps the single most important effect of the reduced surface stress would have been its impact on the oceans. For example there would have been less wind-induced vertical mixing in the ocean, resulting in a shallower, warmer mixed layer, in general, with presumably a smaller overall heat capacity. A smaller heat capacity would make the oceans less able to reduce the amplitude of seasonal and interannual fluctuations of ocean sea surface temperature, which are important in determining the behaviour of the atmosphere²¹. In addition the scales of motion in the ocean would have been reduced by the higher rotation rate of the Earth, which, together with the lower stress, should have weakened the strength of the oceanic gyres. This would also have decreased the amount of heat transported polewards by the oceans, the latter being a significant component of the total polewards flux of energy²². The net effect of all these changes should be to produce a general warming of the tropical oceans and atmosphere and a colder polar region, possibly associated with greater transient departures from mean conditions. The smaller obliquity of the Earth (Fig. 1) might also have complemented these rotationally induced variations.

An additional oceanographic effect to be expected under Precambrian rotational conditions is a reduction in upwelling (depending on continental configurations) and therefore in the supply of nutrients to surface layers. This would have discouraged biological activity and may have some implications to metazoan evolution²³.

All of the above comments should be qualified by noting the tremendous changes which have occurred in oceanic basins during geological history¹³. Some of these changes may well have ameliorated the purely rotational aspects considered here.

A further aspect of the lower wind stress at higher rotation rate concerns atmospheric composition. Transfer of gases between atmosphere and surface would have been restricted, as would, presumably, vertical transport in the troposphere. In addition, the enhanced coldness and aridity at high latitudes would have reduced biological activity, which is crucial to the production of minor, but important, atmospheric constituents such as N_2O , NH_3 and CH_4 . Such subtle effects may have influenced the overall development of atmospheric composition, notably that of O_3 which is considered to be a vital constituent in the development of life on land because of its UV screening efficiency²⁴.

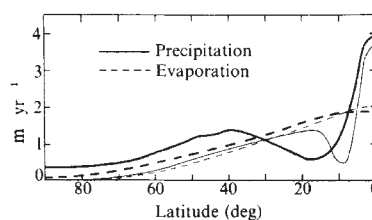


Fig. 5 Comparison of time averaged zonal mean precipitation and evaporation rates for normal rotation (thick lines) and fivefold rotation rates (thin lines).

The net consequence of the reduced polewards heat fluxes suggested by the model for Precambrian times is a colder polar region and an enhanced Equator-to-Pole temperature difference as shown in Fig. 4. A 20 K colder pole is indicated for a rotational period of 9 h, unless the polar cloud cover and thickness were different. The latter may well have been the case given the reduced water vapour content of the colder atmosphere and the smaller-scale systems. Certainly Hart's¹⁶ global mean calculations suggest an overall cloud cover of only $\sim 30\%$ at this time in conjunction with a mean surface temperature of 280 K. This temperature is over 10 K colder than that of the fast rotation rate general circulation model, perhaps indicating that simple comparisons between the two approaches are not meaningful.

Although the model implies that the precipitation rate was very small at high latitudes during the Precambrian era (see Fig. 5), conditions should have been conducive to the accumulation of ice cover, depending on the dispositions of the continent(s)¹⁴. On this basis it can be hypothesised that glacial conditions prevailed throughout the Precambrian, and solely because of rotational considerations! The extent of the ice cover would have declined continuously, all else being equal, as the Earth's rotation rate decreased and the improved circulation systems warmed high latitudes. This explanation can also reasonably account for the termination of the Precambrian ice age 6×10^8 yr ago, when the rotational period of the Earth would have been 20 h and thus dynamical conditions less conducive to the maintenance of glaciation. The occurrence of the Permo-Carboniferous ice age ($\sim 3 \times 10^8$ yr ago) and the present Quaternary ice age are somewhat paradoxical, especially when the assumed increase of the solar output¹⁶ is taken into consideration. Continental drift, volcanic activity and so on may play some part but other unknown factors seem to be important. For example, energy budget studies indicate that in present conditions an ice-free Arctic Ocean would be self-sustaining²⁵, thus raising the problem of the initial cause of the glaciation.

In Fig. 5 the zonal mean evaporation and precipitation rates for the fast and control models are compared. Interpolating results to Precambrian conditions indicates that equatorial conditions were similar to those of today but that the adjacent

zone of high aridity was narrower and closer to the Equator. The general aridity and/or coldness at higher latitudes implied by Figs 4 and 5 suggests a rather harsh climate unattractive to all but the hardiest life-forms at that time. Higher latitudes would have become progressively habitable as the Earth's rotation rate slowed, but the subtropical very arid zone would have expanded. Because of the smaller-scale and less organised nature of meteorological systems suggested for the Precambrian, precipitation generally would have been less intense and more localised. This means that run-off and erosion would have been lower and this could have some relevance to the interpretation of, say, alluvial deposits.

Although the above discussion of Precambrian conditions is necessarily speculative, certain features can be viewed with some confidence. Among these are the delineation of climatic zones in Fig. 5, and the jet stream variations of Fig. 3, both of which are closely associated with the requirements for conservation of absolute angular momentum in the Earth-atmosphere system. Similarly the small scale size of dynamical features is a well established characteristic of high rotation rate^{19,20}. Thus, although distortion of the model results could possibly have arisen from inadequacies of the radiative scheme used, grossly different results should not be expected.

Finally, since a general appreciation of the past variation of the Earth's rotation rate appears to be lacking, it might be useful to reinterpret data from several other disciplines where rotational effects could prove to be of importance.

Received 20 April; accepted 7 August 1979.

1. Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 739 (1976).
2. Lamb, H. H. *Phil. Trans. R. Soc. Lond. A* **266**, 425 (1970).
3. Cox, A. *Met. Monogr.* **8**, 112 (1968).
4. Eddy, J. A. *Climat. Change* **1**, 173 (1977).
5. Wilson, A. T. *Can. J. Earth Sci.* **6**, 911 (1969).
6. Weyl, P. K. *Met. Monogr.* **8**, 37 (1968).
7. Monin, A. S. *Earth's Rotation and Climate*, 1-140 (Radok-Radhakrishna Prakashan, Delhi, 1974).
8. Hunt, B. G. *Mon. Weath. Rev.* **104**, 333 (1976).
9. Hunt, B. G. *J. Atmos. Sci.* **35**, 1133 (1978).
10. Hunt, B. G. *J. Atmos. Sci.* (in the press).
11. Mohr, R. E. in *Growth Rhythms and the History of the Earth's Rotation* (eds Rosenberg, G. D. & Runcorn, S. K.) 43-56 (Wiley, London, 1975).

12. MacDonald, G. J. F. *Rev. Geophys.* **2**, 467 (1964).
13. Irving, E. *Nature* **270**, 304 (1977).
14. Sellers, A. & Meadows, A. J. *Nature* **254**, 44 (1975).
15. Thierstein, H. R. & Berger, W. H. *Nature* **276**, 461 (1978).
16. Hart, M. H. *Icarus* **33**, 23 (1978).
17. Wetherald, R. T. & Manabe, S. *J. Atmos. Sci.* **32**, 2044 (1975).
18. Manabe, S., Smagorinsky, J., Holloway, J. L. & Stone, H. M. *Mon. Weath. Rev.* **98**, 175 (1970).
19. Hide, R. *Q. J. R. met. Soc.* **103**, 1 (1977).
20. Fultz, D. et al. *Met. Monogr.* **4** (1959).
21. Namias, J. *Mon. Weath. Rev.* **106**, 1557 (1978).
22. Oort, A. H. & von der Haar, T. H. *J. phys. Ocean.* **6**, 781 (1976).
23. Knoll, A. H. *Nature* **275**, 701 (1978).
24. Berkner, L. V. & Marshall, L. C. *Discuss. Faraday Soc.* **37**, 122 (1964).
25. Donn, W. L. & Shaw, D. M. *J. geophys. Res.* **71**, 1087 (1966).

Single strands induce recA protein to unwind duplex DNA for homologous pairing

Richard P. Cunningham, Takehiko Shibata, Chanchal DasGupta & Charles M. Radding

The Departments of Internal Medicine, and Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06510

Single-stranded DNA, whether homologous or not, stimulates purified Escherichia coli recA protein to unwind duplex DNA. This helps to explain how recA promotes a search for homology in genetic recombination. As oligodeoxynucleotides also stimulate unwinding, a common mechanism may relate the function of recA protein in recombination to other functions (SOS) induced by oligonucleotides.

BECAUSE breakage of DNA by various means stimulates general genetic recombination, some investigators have sugges-

ted that recombination begins with the interaction of a single strand with duplex DNA (see ref. 1 for review). As a model for this kind of interaction, Holloman *et al.*² and Beattie *et al.*³ studied the uptake of homologous single strands by superhelical DNA [replicative form I (RFI) or form I], a reaction that produces a D-loop (see Fig. 1). By transfection of *Escherichia coli*, Holloman and C.M.R.⁵ implicated the *recA* gene in the homologous interaction of a single-stranded fragment with superhelical DNA. More recently, Shibata *et al.*⁴ observed that stoichiometric amounts of highly purified *recA* protein promoted the rapid formation of D-loops by superhelical DNA and homologous single-stranded fragments (Fig. 1).

The properties of the uncatalysed formation of D-loops, which occurred rapidly only at non-physiological temperatures,

suggested that the rate was limited by the initial unstacking of a small number of base pairs in superhelical DNA³. As recA protein bypassed this rate-limiting step, we suggested that it acted in part by unstacking base pairs in duplex DNA. We also observed that a preparation of nicked circular DNA (form II) was one-third as active as superhelical DNA in forming complexes with single-stranded fragments⁴. Subsequently, we have confirmed that form II DNA is a substrate for the formation of D-loops by recA protein (see below). The latter observation strengthened the inference that recA protein has an unstacking or unwinding activity, which is the subject of this article.

Unwinding of duplex DNA by recA protein in the presence of ATP γ S and single strands

To obtain direct evidence that recA protein unwinds DNA, we incubated it with fd form II DNA, added DNA ligase, and examined the product by gel electrophoresis and isopycnic centrifugation in CsCl plus propidium diiodide^{6,7}.

In the absence of recA protein, ligase sealed more than half of the form II DNA as indicated by gel electrophoresis (Fig. 2a, b) and by the assay of Kuhnlein *et al.*⁸ for covalently closed DNA. At the ionic strength of the gel, these covalently closed molecules (form I') migrated as a set of bands of positively superhelical DNA, which overlapped a band due to linear DNA (form III) present in the preparation (Fig. 2b). When recA protein and adenosine 5'-O-(3-thiotriphosphate) (ATP γ S) were incubated

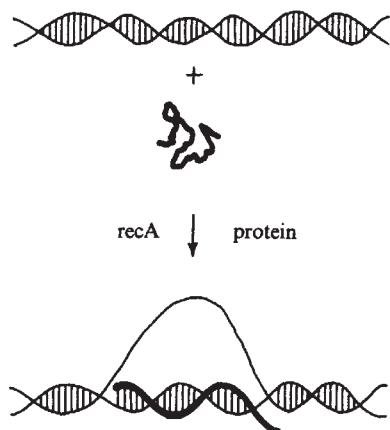


Fig. 1 The formation of a D-loop catalysed by recA protein⁴. The actions of recA protein apparently include (1) unwinding double-stranded DNA, (2) unfolding single-stranded DNA, (3) homologous pairing, and (4) uptake of a single strand rather than rewinding of the original duplex DNA.

with form II DNA before ligation, the observed distribution of products was similar; ligase worked well but there was no apparent change in the average linking number of the population of sealed molecules (Fig. 2c). However, when we added fragments of single-stranded Φ X174 DNA during the incubation with recA protein, we observed a very different result. Ligase still closed about one-third to one-half of the fd form II DNA, but all the closed DNA migrated as a single band of more compact DNA, at about the position of form I (Fig. 2d). In three experiments, when recA protein and ATP γ S were incubated with form II DNA in the absence of any added single-stranded DNA before ligation, we observed the formation of a small amount of material with the electrophoretic mobility of form I (see Table 1, lines 1 and 2). Controls included the following variations of the mixture represented in Fig. 2d: no recA protein; no form II DNA; no ATP γ S; ATP in place of ATP γ S;

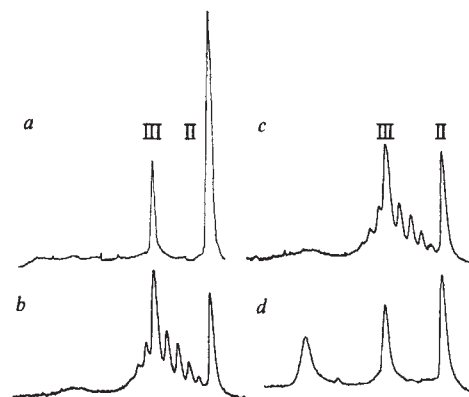


Fig. 2 The linking number of circular DNA sealed in the presence of recA protein and ATP γ S (Boehringer-Mannheim). Microdensitometer tracings of electrophoretic gels prepared as described in Fig. 4, legend. a, Control, fd form II DNA incubated with recA protein but no ligase. The preparation of DNA contained about 30% form III DNA as indicated. b, Ligase only (New England BioLabs). c, Ligase plus recA protein. d, Ligase, recA protein and 12- μ M fragments of heterologous single-stranded DNA from Φ X174. The various forms of circular DNA from phages Φ X174 and fd were prepared as cited previously⁴. Single-stranded fragments were prepared by boiling 840 μ M DNA for 15 min in 0.1 mM EDTA and 10 mM Tris-HCl at pH 7.5. All concentrations of DNA are expressed as mol of nucleotide. The reaction mixture was the same as that described in Table 2 legend, except that heterologous single-stranded fragments were substituted for homologous fragments, the volume was doubled and ATP was replaced by 0.5 mM ATP γ S. After incubation at 37 °C for 30 min, the mixture was chilled on ice, KCl, (NH₄)₂SO₄ and NAD⁺ added to make their concentrations 25 mM, 10 mM, and 25 μ M, respectively, and incubated at 18 °C for 5 min. 0.3 units of *E. coli* DNA ligase were then added and the sample, 50 μ l total volume, incubated for 90 min. The reaction was stopped by adding EDTA, 28 mM, and SDS, 0.17%. The sample was transferred to a bath at 37 °C, proteinase K, 0.2 mg ml⁻¹ added and the sample incubated for 15 more minutes.

recA protein in the absence of ligase (Fig. 2a). In no case was DNA detected in the position of the rapidly migrating band.

To determine whether the compact closed DNA had a lower or a higher linking number than the DNA closed by ligase alone (form I'), we centrifuged the products of a similar experiment in an isopycnic gradient of CsCl and propidium diiodide (Fig. 3 and Table 1). In the absence of recA protein, half of the recovered DNA was found in a single peak at the dense end of the gradient, as expected for form I' DNA derived from the closure of form II by ligase (Fig. 3b, Table 1, line 1). Incubation of form II DNA with recA protein and ATP γ S before ligation had little effect on the distribution (Fig. 3a, Table 1, line 2). However, when recA protein, fd form II DNA, Φ X174 single-stranded fragments and ATP γ S were incubated before ligation, the peak of form I' DNA at the dense end of the gradient disappeared almost entirely and was replaced by DNA with a broad distribution of densities

Table 1 Unwinding of DNA

	Total recovery (%)	Distribution (%)		
		I'	Intermediate	II + III
Ligase only	22	46	4	50
+recA protein	29	39	7	54
+recA protein and fragments	28	8	21	71

Data tabulated from Fig. 3. The DNA labelled 'Intermediate' has a lower linking number than form I' DNA, the product of sealing by ligase alone.

between those of form I' and linear or form II DNA (Fig. 3a, Table 1, line 3), that is, DNA with lower linking numbers than form I'. The peak of material of intermediate density corresponded to the position of negatively superhelical fd DNA (form I DNA). The total recovery of radioactivity was similar in the three parts of this experiment; the fraction of closed circular DNA relative to linear or form II DNA (Fig. 3a, Table 1, line 1) was the same as that estimated by the assay of Kuhnlein *et al.*⁸, which indicates that the distribution of recovered material was not biased.

Lack of topoisomerase activity of recA protein

The unwinding observed in the experiments described above could have occurred before or after ligation, but three experiments show that recA protein lacks detectable topoisomerase activity in the presence of either ATP γ S or ATP. We incubated recA protein with relaxed closed circular DNA (form I') in

conditions identical to those of the experiments shown in Figs 2 and 3, including a sample which contained heterologous fragments and ATP γ S. When we observed the product by gel electrophoresis, we found no DNA in the position of form I DNA, and no perceptible shift in the distribution of form I' DNA (Fig. 4). With the following additions or omissions, we examined, by gel electrophoresis, the fate of fd RFI incubated for 30 min at 37 °C with recA protein and fragments of single-stranded DNA: ATP; ATP γ S; ATP γ S without single-stranded fragments; ATP γ S without recA protein. We observed little or no change in the amount of form I DNA (data not shown).

In further experiments we made D-loops by incubating recA protein with fd form I DNA and fragments of fd single strands in

Table 2 Homologous pairing of double-stranded DNA and single-stranded fragments

Fragments	% fd double-stranded [³ H]DNA retained	
	form I	form II
None	1	4
fd	34	22
Φ X	2	6

The reaction mixture, 20.5 μ l in volume, contained 31 mM Tris-HCl at pH 7.5, 25 mM MgCl₂, 1.8 mM dithiothreitol, 88 μ g ml⁻¹ bovine serum albumin, 1.3 mM ATP, 8.8 μ M RF[³H]DNA, 12 μ M single-stranded DNA fragments and 98 μ g ml⁻¹ recA protein⁴. The mixture was incubated at 37 °C for 30 min, chilled and 18 μ l transferred to 0.5 ml of cold 25 mM EDTA. Aliquots were taken from this mixture to measure total counts, D-loops and closed circular DNA, exactly as described before⁴.

the presence of ATP⁴. Measurement by electron microscopy of 22 D-loops with one or two single-stranded tails showed that the mean length was 416 $\pm$ 45 nucleotides. (We were unable to distinguish the double-stranded arm of the D-loop from the single-stranded one. As there was little difference in the lengths of the two arms of any D-loop, we simply measured both and calculated the average of all values.) In the conditions of the reaction, fd form I DNA should contain about 40 superhelical turns which corresponds to a D-loop about 400 nucleotides long. Previous observations have shown a rough equivalence between the length of D-loops measured by electron microscopy and the length predicted from the superhelix density in solution⁹. Observation of the same relationship in these experiments suggests that when recA protein catalysed the formation of D-loops from form I DNA, the linking number of the closed circular DNA was not markedly changed.

Nature of the cofactor for unwinding by recA protein

Unwinding is supported either by heterologous or homologous single-stranded fragments (Fig. 4a, b). Circular single-stranded DNA, which does not form D-loops, promoted unwinding (Fig. 4d). Double-stranded restriction fragments were inactive, but their boiled products were active (Fig. 4f, g). This observation indicates that the cofactor was not some fortuitous ingredient of the DNA preparations. A digest, consisting entirely of acid-soluble material produced by treatment with pancreatic DNase and boiling, also supported the unwinding activity very well (Fig. 4e). The following, added singly, promoted little or no unwinding: deoxyadenosine, dAMP, dGMP, dCMP, dTMP, AMP and CMP (data not shown).

Homologous pairing of form II DNA and single-stranded fragments

Table 2 compares form II and form I DNA as substrates for recA protein in the homologous pairing with single-stranded

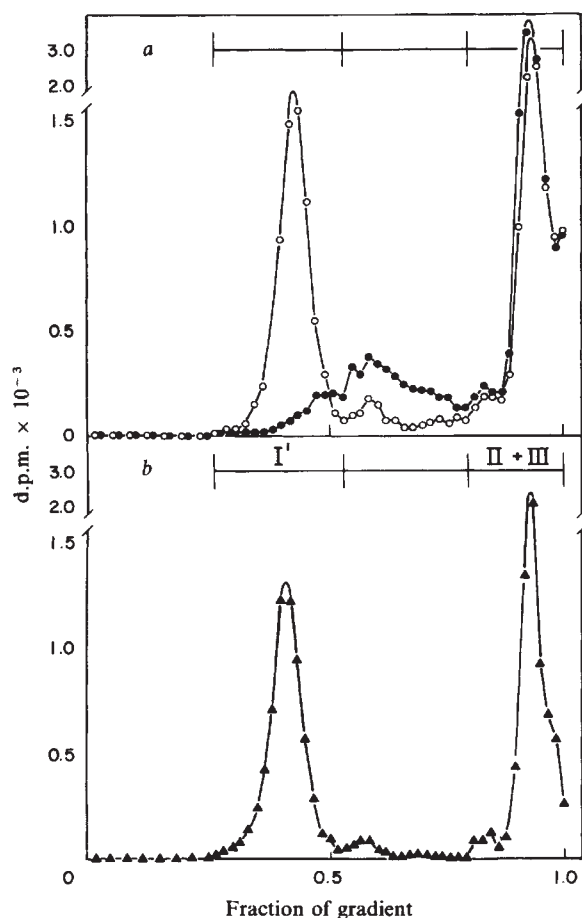


Fig. 3 Decreased linking number of circular DNA sealed in the presence of recA protein and ATP γ S; isopycnic centrifugation in CsCl and propidium diiodide⁷. a, fd form II DNA sealed in the presence of recA protein only (○), and in the presence of recA protein plus heterologous fragments of single-stranded DNA from Φ X174 (●). b, Control. DNA sealed in the absence of recA protein (▲). The left end of the figure corresponds to the dense end of the gradient. We summed values as indicated by the brackets (Table 1). All values were corrected for quenching. Solutions, 7.4 ml, contained CsCl to make the density 1.565 g cm⁻³, 500 μ g ml⁻¹ propidium diiodide, 10 mM Tris-HCl and 1 mM EDTA at pH 8.0. These solutions were centrifuged at 44,000 r.p.m. and 20 °C for 69 h in a Beckman 75ti rotor. The resulting gradients were collected from the bottom of the tube and counted in a scintillation fluor containing Triton X-100.

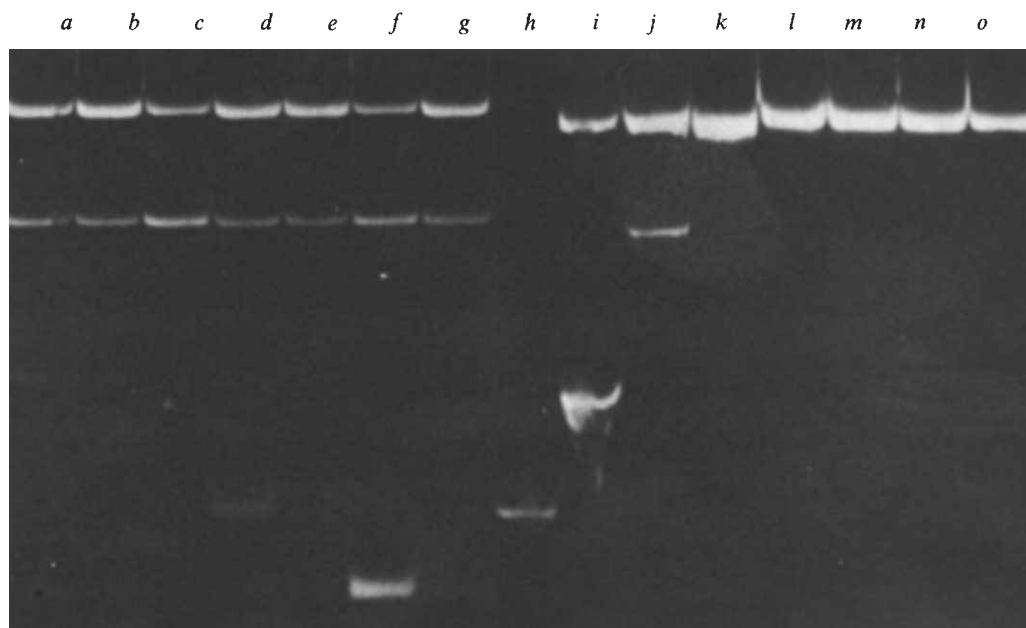


Fig. 4 Nature of the cofactor that stimulates unwinding by *recA* protein in the presence of $\text{ATP}\gamma\text{S}$ (*a-g*); absence of topoisomerase activity in *recA* protein (*l-o*): agarose gel electrophoresis. Lanes *a-g* contained fd form II DNA sealed in the presence of *recA* protein, $\text{ATP}\gamma\text{S}$ and $12\ \mu\text{M}$ DNA cofactors: *a*, fragments of single-stranded ΦX174 DNA; *b* fragments of single-stranded fd DNA; *c*, no addition; *d*, fd phage DNA; *e*, boiled DNase I limit digest of ΦX174 RFI DNA; *f*, restriction endonuclease *Hae*III digest of fd RFI; *g*, the preceding digest boiled. Lanes *h-k* contained markers: *h*, fd phage DNA; *i*, fd RFI DNA; *j*, fd form II DNA ($\sim 70\%$) and fd form III DNA ($\sim 30\%$), *k*, fd form I' DNA. Lanes *l-o* contained fd form I' DNA incubated with *recA* protein and $\text{ATP}\gamma\text{S}$: *l*, no addition; *m*, fragments of single-stranded ΦX174 DNA; *n*, preceding sample minus *recA* protein; *o*, fragments of single-stranded ΦX174 DNA plus ATP instead of $\text{ATP}\gamma\text{S}$. To obtain the limit digest indicated above $24\ \mu\text{M}$ ΦX174 RF DNA was treated with $20\ \mu\text{g ml}^{-1}$ pancreatic DNase (Worthington) for 30 min at 37°C in $5\ \text{mM}$ MgCl_2 and $10\ \text{mM}$ Tris-HCl at $\text{pH } 7.5$, which rendered all the DNA acid soluble. The product was boiled for 4 min to inactivate the pancreatic DNase and denature all the fragments. Nicked circular DNA, sealed by polynucleotide ligase was slightly positively superhelical under the conditions of gel electrophoresis (for example lane *c*) whereas closed circular DNA relaxed by calf thymus topoisomerase had a titrable superhelix density close to zero (for example lane *k*). We used 1% agarose gels, $16\ \text{cm} \times 16\ \text{cm} \times 0.3\ \text{cm}$ in E buffer, which contained $40\ \text{mM}$ Tris-acetate, $5\ \text{mM}$ sodium acetate and $1\ \text{mM}$ EDTA at $\text{pH } 8.0$. Current was applied at $35\ \text{V}$ for 16–18 h, and the buffer was recirculated between the reservoirs, following which we stained the gels for 2 h in E buffer containing $0.5\ \mu\text{g ml}^{-1}$ ethidium bromide. We illuminated the gels from below with a short-wavelength UV lamp (Ultra-Violet Products) and photographed them with Polaroid Tyffe 55 Land film. The negatives were scanned with a Joyce-Loebl recording micro-densitometer.

fragments. In the conditions of these experiments (see Table 2, legend), which differ slightly from previously described conditions⁴, the preparation of form II DNA was one-half to two-thirds as active as form I DNA in making complexes that were retained by nitrocellulose filters. As the preparation of form II DNA contained less than a few per cent of form I DNA, these complexes must have been made with DNA that was not superhelical. The synthesis of complexes with form II DNA required homologous single-stranded fragments, and the properties of the product were similar to those of D-loops made from RFI; the complexes were not dissociated by treatment with 0.2% SDS nor by heating at 50°C for 4 min in $1.5\ \text{M}$ NaCl and $0.15\ \text{M}$ Na citrate at $\text{pH } 7$. Preliminary electron microscopic observations suggest that a significant fraction of complexes made with form II DNA contain D-loops (data not shown). Recently, McEntee, Weinstock and Lehman also described the synthesis of D-loops from non-superhelical DNA, in that case, from linear phage P22 DNA¹⁰.

The formation of D-loops by form II DNA suggests that *recA* protein unwinds DNA in the presence of ATP as well as in the presence of the analogue, $\text{ATP}\gamma\text{S}$.

Discussion

Other experiments have shown that $\text{ATP}\gamma\text{S}$ competitively inhibits the ATPase activity of *recA* protein and blocks the formation of D-loops, but promotes the formation of complexes of protein, double-stranded DNA and single-stranded DNA

(unpublished observations). The competitive inhibition shows that $\text{ATP}\gamma\text{S}$ occupies the same binding site as ATP; this, together with the other observations cited, suggests that the accumulation of partially unwound molecules reflects a step in the reaction that normally produces a D-loop. These effects of $\text{ATP}\gamma\text{S}$ on *recA* protein are reminiscent of observations on DNA gyrase which also catalyses several partial reactions in the presence of another non-hydrolysable analogue of ATP¹¹. From the observed synthesis of D-loops with form II DNA, we infer that *recA* protein also unwinds DNA in the presence of ATP. The hydrolysis of ATP alters or dissociates the protein-DNA complex and accounts for the failure of partially unwound molecules to accumulate in the presence of ATP and DNA ligase (unpublished observations).

At first glance, the activation of unwinding by heterologous single strands seems incongruous for a mechanism that promotes homologous pairing, but further thought suggests that the opposite is true, that nonspecific stimulation of unwinding by any single strand explains in part how *recA* protein promotes a search for homology. As recombination in *E. coli* depends strongly on the function of *recA*¹², the action of *recA* protein by this mechanism implies that recombination often begins with the interaction of a single strand with duplex DNA.

Mutants of *recA* are pleiotropic, they affect not only recombination, but also a set of inducible functions, the so-called SOS functions¹³; these include error-prone repair, inactivation of phage repressors and the synthesis of *recA* protein itself (for reviews see refs 14, 15). Damaged DNA or blocked replication

are inducing stimuli, and oligonucleotides have been implicated as intermediates in the induction^{16,17}. The *tif-1* mutation, which is located in the *recA* structural gene, makes the synthesis of *recA* protein thermo-inducible¹⁸⁻²², indicating that a suitable form of the protein stimulates its own synthesis. The interaction of oligonucleotides with wild-type protein presumably stimulates synthesis by the same mechanism. However, observations on the *tsI* mutation indicate that abundant synthesis of *recA* protein may not be sufficient to induce fully all of the SOS functions^{19,23}. Do oligonucleotides have a direct role in the inducible functions? Roberts *et al.* have studied one of the SOS functions *in vitro*, namely the inactivation of λ repressor by proteolytic cleavage^{24,25}. In spite of the difficulty of reconciling proteolysis and unwinding of DNA, there are marked similarities between our observations on the unwinding of DNA by wild-type *recA* protein and those of Roberts *et al.* on inactivation of λ repressor by *tif recA* protein. Both processes require ATP or ATP γ S, and oligonucleotide or single-stranded polynucleotide. The experiments described in this article show that both oligonucleotides and single-stranded DNA affect the functional properties of *recA* protein. The stimulation of

unwinding by oligonucleotides, as well as by single strands, suggests that a common mechanism relates the function of *recA* protein in recombination to the other functions that are induced by oligonucleotides. Accordingly, oligonucleotides or single strands may trigger a conformational change of *recA* protein that can both unwind DNA and alter the interaction of *recA* protein with certain other proteins that bind to DNA.

A further inference from the similar effect of single strands and oligonucleotides on *recA* protein is that recombination may be inducible to some extent. Induction of synthesis of *recA* protein by single strands would provide the requisite amounts of protein to form D-loops, amounts which are large *in vitro*⁴. On the other hand, the induction of *recA* protein synthesis by oligonucleotides would not necessarily produce an increase in the number of genetic exchanges if any other factor, such as single strands, were limiting.

T.S. is a visiting fellow from The Institute of Physical and Chemical Research, Saitama, Japan. We acknowledge the technical assistance of Lynn Osber. This research was sponsored by grant no. NP 90E from the American Cancer Society and grant no. PHS CA 16038-05 from the NCI.

Received 4 May; accepted 18 June 1979.

1. Radding, C. M. A. *Rev. Biochem.* **47**, 847-880 (1978).
2. Holloman, W. K., Wiegand, R., Hoessli, C. & Radding, C. M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2394-2398 (1975).
3. Beattie, K. L., Wiegand, R. C. & Radding, C. M. *J. molec. Biol.* **116**, 783-803 (1977).
4. Shibata, T., DasGupta, C., Cunningham, R. P. & Radding, C. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1638-1642 (1979).
5. Holloman, W. K. & Radding, C. M. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3910-3914 (1976).
6. Liu, L. F. & Wang, J. C. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2098-2102 (1978).
7. Anderson, P. & Bauer, W. *Biochemistry* **17**, 594-601 (1978).
8. Kuhnlein, U., Penhoet, E. E. & Linn, S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1169-1173 (1976).
9. Wiegand, R. C., Beattie, K. L., Holloman, W. K. & Radding, C. M. *J. molec. Biol.* **116**, 805-824 (1977).
10. McEntee, K., Weinstock, G. M. & Lehman, I. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2615-2619 (1979).
11. Sugino, A., Higgins, N. P., Brown, P. O., Peebles, C. L. & Cozzarelli, N. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4838-4842 (1978).

12. Clark, A. J. A. *Rev. Genet.* **7**, 67-86 (1973).
13. Radman, M. in *Molecular Mechanisms for Repair of DNA Part A* (eds Hanawalt, P. C. & Setlow, R. B.) 355-367 (Plenum, New York, 1975).
14. Devoret, R. *Biochimie* **60**, 1135-1140 (1978).
15. Witkin, E. M. *Bact. Rev.* **40**, 869-907 (1976).
16. Oishi, M. & Smith, C. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3569-3573 (1978).
17. Oishi, M., Smith, C. L. & Friefeld, B. *Cold Spring Harb. Symp. quant. Biol.* **43**, 897-907 (1979).
18. Castellazzi, M., Morand, P., George, J. & Buttin, G. *Molec. gen. Genet.* **153**, 297-310 (1977).
19. McEntee, K. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5275-5279 (1977).
20. Gudas, L. J. & Mount, D. W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5280-5284 (1977).
21. Little, J. W. & Kleid, D. G. *J. biol. Chem.* **252**, 6251-6252 (1977).
22. Emmerson, P. T. & West, S. C. *Molec. gen. Genet.* **155**, 77-85 (1977).
23. Mount, D. W., Kosel, C. K. & Walker, A. *Molec. gen. Genet.* **146**, 37-41 (1976).
24. Roberts, J. W., Roberts, C. W. & Craig, N. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4714-4718 (1978).
25. Roberts, J., Roberts, C. W., Craig, N. L. & Phizicky, E. *Cold Spring Harb. Symp. quant. Biol.* **43**, 917-920 (1979).

The *src* gene product of transformed and morphologically reverted ASV-infected mammalian cells

Marc S. Collett, Joan S. Brugge & Raymond L. Erikson

Department of Pathology, University of Colorado Medical Center, 4200 East Ninth Avenue, Denver, Colorado 80262

Alan F. Lau, Richard A. Krzyzek & Anthony J. Faras*

Department of Microbiology, University of Minnesota Medical School, 925 Mayo Memorial Building, Minneapolis, Minnesota 55455

Morphological revertants of avian sarcoma virus transformed vole cells contain the sarcoma gene product (pp60^{src}) in an enzymatically active form, suggesting that the presence of pp60^{src} protein kinase activity is insufficient to induce morphological transformation. Structural analyses of pp60^{src} from infected vole cell clones suggest that in one of the revertant clones an alteration in pp60^{src} may be responsible for morphological reversion while in a second clone, reversion may result from an alteration in a cell gene product with which pp60^{src} must interact. As these morphological revertant cells are tumorigenic, different cell components are required to interact with pp60^{src} to facilitate the two events.

AVIAN SARCOMA VIRUSES (ASV) are able to induce tumours in birds and a variety of mammals as well as stably transform both avian and mammalian cells in culture. A single ASV gene, designated *src*, has been shown to be responsible for the induction and maintenance of cell transformation *in vitro* and tumour production in infected animals^{1,2}. The ASV transforming or *src* gene product is a 60,000-molecular weight (MW) phosphoprotein (pp60^{src})³⁻⁶ with an associated protein kinase activity, suggesting that ASV-induced neoplastic transformation may be the result of aberrant phosphorylation of cellular proteins by pp60^{src} (refs 6-8). In an attempt to elucidate the mechanism(s) by which the expression of this particular viral gene is regulated, and hopefully ultimately provide information concerning virus-induced oncogenesis, we have compared the products of the *src* gene in virus-transformed mammalian cells and morphologically normal revertant subclones of these cells. The cell system chosen for these studies was ASV-transformed

* To whom reprint requests should be addressed.

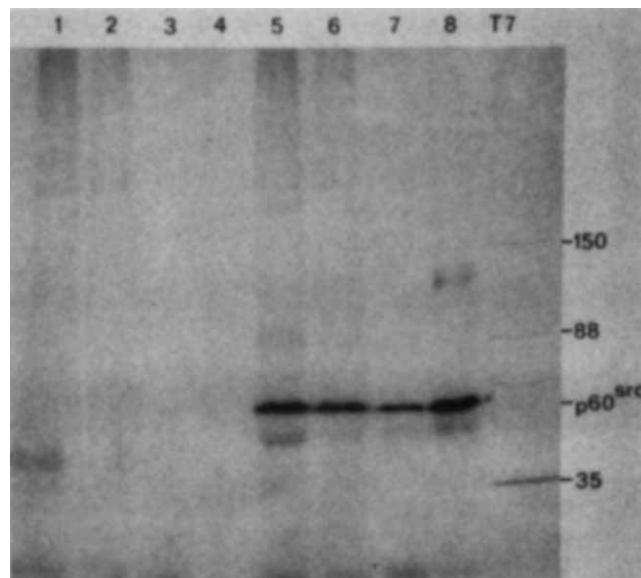


Fig. 1 Presence of phosphorylated pp60^{src} in transformed and revertant vole cells. Cultures were radiolabelled for 2 h with ³²P-orthophosphate (carrier-free (ICN), 0.5 mCi ml⁻¹). After cell lysis and clarification, cell extracts were immunoprecipitated with either normal rabbit serum or TBR serum as previously described^{3,7}. Portions of the immunoprecipitated proteins were subjected to discontinuous SDS-polyacrylamide gradient gel electrophoresis (5–15% acrylamide)³, and the fixed, dried gels were autoradiographed. Tracks 1–4, normal rabbit serum; tracks 5–8, TBR serum. Tracks 1, 5, transformed clone 1T; tracks 2, 6, transformed clone 22T; tracks 3, 7, revertant subclone 866R; tracks 4, 8, revertant subclone 4R. ³⁵S-methionine-labelled T7 virion proteins were included as molecular weight markers; numbers in kilodaltons.

European field vole (*Microtus agrestis*) cells, since not only could we obtain revertants of these transformed cells that exhibited the normal cellular morphology, but also these particular cells contained levels of pp60^{src} similar to the transformed vole cells^{9–12}. Moreover, as fully infectious transforming ASV can be recovered from the revertant vole subclones, loss of the transformed phenotype did not seem to result from a mutation affecting the *src* gene sequence^{9–11}. Thus the reversion phenomenon in this vole cell system appeared to reflect either some post-translational alteration or lack thereof in pp60^{src}, or a defect in a host cell component(s) with which the *src* gene product must interact.

We report here that, although a slight alteration in the primary structure of pp60^{src} obtained from revertant vole cells was detected, this alteration does not seem to be responsible for loss of morphological transformation in the revertant vole cell types, as the slightly altered pp60^{src} molecule was also found in chick cells morphologically transformed by RSV rescued from these revertant vole cell subclones as well as subclones of spontaneous retransformants. Biochemical analysis of pp60^{src} from two revertant vole cell subclones indicates that, in at least one of the clones, morphological reversion is the result of an alteration in a cellular function rather than *src* itself. pp60^{src} isolated from these revertant vole cells exhibits protein kinase activity, suggesting that the presence of such an activity in cells is itself insufficient to induce morphological transformation. More importantly, as both of the revertant vole cells are tumorigenic it seems that one of the phenotypic parameters of cellular transformation can be expressed in the absence of another.

Phosphorylated pp60^{src} in transformed and revertant cells

Once the presence of the *src* gene product in the revertant vole cell clones¹² was established our initial studies were directed at

determining whether any subtle differences could be detected in the structure of pp60^{src} obtained from transformed and revertant vole cells. Because the *src* gene product is a phosphoprotein, studies were carried out to determine if differences in phosphorylation existed between pp60^{src} obtained from transformed and revertant vole cells. ASV (Schmidt-Ruppin strain, subgroup D) transformed and morphologically reverted vole cells were radiolabelled with ³²P-orthophosphate and cell extracts were prepared by previously described procedures^{3,7}. After immunoprecipitation with either non-immune rabbit serum or serum from rabbits bearing ASV-induced tumours (TBR serum)³, the radiolabelled materials were subjected to electrophoresis in SDS-polyacrylamide gels. TBR serum precipitated a phosphorylated 60,000-MW protein from both transformed and both revertant vole cell lines (Fig. 1), whereas no such protein could be observed when non-immune rabbit serum was used for immunoprecipitation or when uninfected

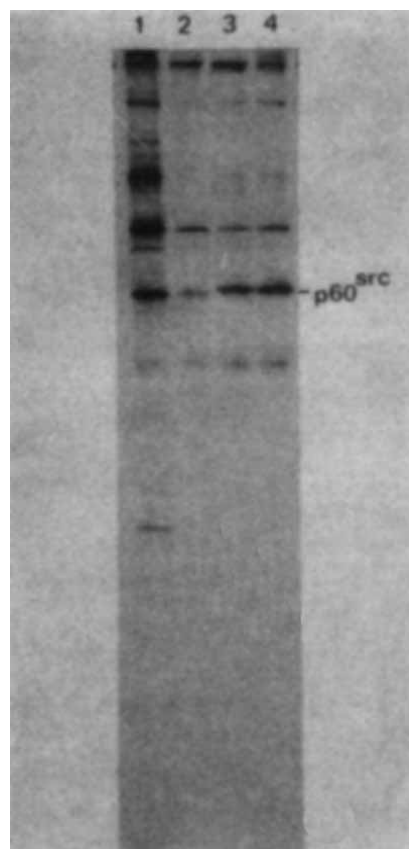
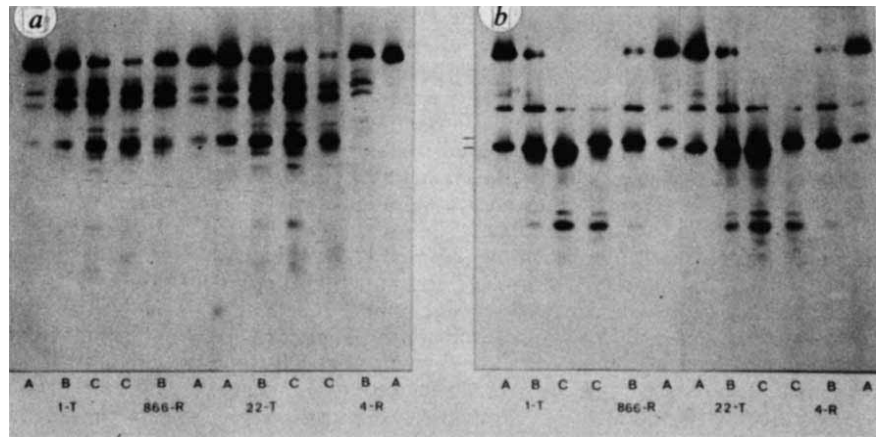


Fig. 2 Detection of a difference in electrophoretic mobility in SDS-containing polyacrylamide gels of pp60^{src} from transformed and revertant vole cells. ³⁵S-methionine labelled samples immunoprecipitated with TBR serum as described in Fig. 1 were electrophoresed in a discontinuous SDS-containing polyacrylamide slab gel consisting of a 4.2% acrylamide, 0.1% bisacrylamide stacking gel and a 10% polyacrylamide (38:1 acrylamide: bisacrylamide) separation gel, 14 cm long. Track 1, transformed clone 1T; track 2, transformed clone 22T; track 3, revertant subclone 866R; track 4, revertant subclone 4R.

vole cell extracts were immunoprecipitated with this TBR serum. The intensity of the pp60^{src} bands in the autoradiogram suggests that the *src* gene product in both transformed and revertant cells was phosphorylated to a similar extent. Preliminary results have indicated that proteolytic peptides of phosphorylated pp60^{src} from transformed and revertant cell lines are similar. However, more detailed studies concerning the phosphorylation of the *src* protein in this system are now in progress.

Fig. 3 Partial proteolysis mapping of pp60^{src} from transformed and revertant vole cells. ³⁵S-methionine labelled pp60^{src} proteins immunoprecipitated from the various vole cell lines were excised from preparative polyacrylamide gels and subjected to limited proteolysis during re-electrophoresis as described previously¹³. *a*, Digestion with chymotrypsin. Tracks A, 0.1 µg enzyme; tracks B, 0.6 µg enzyme; tracks C, 2.5 µg enzyme. *b*, Digestion with *Staphylococcus aureus* V8 protease. Tracks A, 0.001 µg enzyme; tracks B, 0.005 µg enzyme; tracks C, 0.05 µg enzyme. The lines in the margin of *b* indicate the V8 protease 26,000-MW peptide (the C-terminal 40% of the pp60^{src} protein molecule)¹⁴ which shows an altered mobility between the revertant and transformed cell proteins.



Difference in electrophoretic mobility of pp60^{src} from transformed and revertant cells

³⁵S-methionine labelled proteins were also immunoprecipitated from extracts of both clones of transformed and revertant vole cells and electrophoresed on a slab gel longer than used in our initial studies in an effort to better resolve the pp60^{src} proteins (Fig. 2). The pp60^{src} proteins isolated from the two transformed cell lines co-migrated, as did the proteins from the two revertant subclones. However, the proteins from the revertant cell lines showed a slightly slower electrophoretic mobility than those from the transformed clones. As the pp60^{src} proteins from transformed vole and chicken cells co-migrate in these conditions (data not shown), these results suggest that the pp60^{src} protein present in the two revertant clones may represent an altered *src* gene product.

Proteolysis mapping of pp60^{src}

For a better assessment of the similarities and any possible differences in the pp60^{src} proteins from the transformed and revertant vole cells, one-dimensional peptide analyses were performed on ³⁵S-labelled pp60^{src} with both chymotrypsin (Fig. 3*a*) and *Staphylococcus aureus* V8 protease (Fig. 3*b*)¹³. The chymotryptic peptides generated by this partial proteolysis of the four proteins were all very similar if not identical (Fig. 3*a*). The V8 protease digestion patterns of pp60^{src} from the two transformed cell lines also appeared identical (Fig. 3*b*). However, both revertant cell pp60^{src} proteins exhibited V8 protease peptides that differed in electrophoretic mobility from the corresponding peptides of the two transformed vole cell lines (Fig. 3*b*, margin lines). The initial digestion products resulting from V8 protease digestion of pp60^{src} are two peptides of MW ~34,000 and 26,000 (ref. 14). The larger V8 protease fragment represents the amino-terminal 60% of the pp60^{src} molecule and the smaller fragment encompasses the carboxy-terminal 40% (ref. 14). Thus, it seems that the pp60^{src} C-terminal V8 protease peptide in both of the revertant cell lines was changed to a peptide of slightly slower electrophoretic mobility, confirming the slight alteration in pp60^{src} obtained from these revertant cells observed by polyacrylamide gel electrophoresis (Fig. 2).

In an attempt to determine if the altered V8 protease peptide of pp60^{src} from the two revertant subclones represented either a stable genetic change, or a variation in secondary protein modification or processing by the revertant cells, virus was rescued from the two revertant subclones and used to infect chicken cells. The ³⁵S-methionine-labelled pp60^{src} proteins were then isolated from these rescued virus-infected (and transformed) avian cells and subjected to partial proteolysis mapping using V8 protease as described above. When the protein obtained from the revertant subclone 4R rescued virus-infected chick cells (4R chick) was analysed, peptide fragments identical

to those of pp60^{src} from the 4R revertant vole cells (4R vole) were obtained (Fig. 4*a*), suggesting that the alteration in the C-terminal V8 protease peptide of the *src* protein in this revertant subclone represented a stable change. In addition, we have been able to isolate clones of cells from cultures of our revertant cells grown to high cell density that appear to have regained the transformed phenotype. These re-transformed (RT) clones have also been used for *src* protein analysis. V8 protease peptide mapping of pp60^{src} from a re-transformed clone of the 4R revertant cells (4RT vole) again revealed no change from the pattern observed in the protein from the parent revertant cell line (Fig. 4*a*).

Similar peptide analyses were performed on the *src* protein isolated from chick cells transformed with revertant subclone 866R rescued virus (866R chick) and a clone of re-transformed

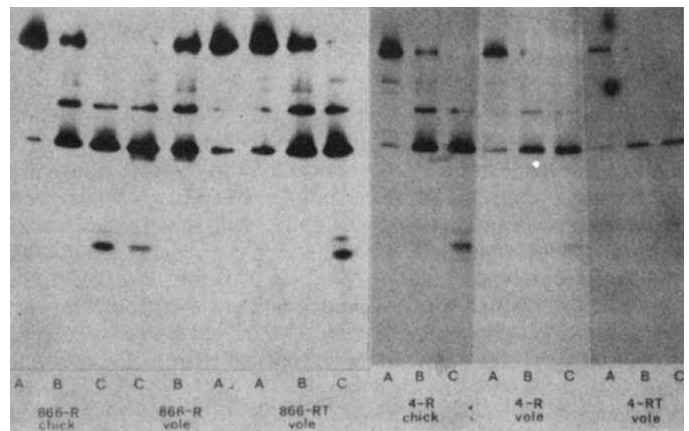


Fig. 4 Partial proteolysis mapping of pp60^{src} from rescued virus-infected chick cells and re-transformed vole cells. The rescue of ASV from the revertant vole cell subclones 4R and 866R by Sendai virus-mediated fusion with chick fibroblasts was performed by the procedure of Boettiger¹⁹ as previously described⁹. The rescued virus (SR-ASV, subgroup D) obtained from the supernatant fluids of the fusion culture was used to infect chick cells. The infection resulted in virus production and the complete transformation of the chick cell culture. Re-transformed cell lines (4RT and 866RT) were obtained by picking cells that appeared morphologically transformed from 4R and 866R revertant subclones that were grown and maintained at high cell densities. The re-transformed cell lines maintain their transformed morphology on continuous subculturing. ³⁵S-methionine labelled pp60^{src} was immunoprecipitated from the rescued virus-infected chick cells (4R chick and 866R chick), the re-transformed vole cells (4RT and 866RT vole), and the original revertant subclones (4R vole and 866R vole) and subjected to V8 protease partial proteolysis as described in Fig. 3 legend. Tracks A, 0.001 µg enzyme; tracks B, 0.005 µg enzyme; tracks C, 0.05 µg enzyme.

vole cells from this revertant cell line (866RT vole) (Fig. 4b). As was the case with the 4R revertant cells the two major V8 protease peptides of pp60^{src} from these cell types were all identical. Although it first appeared that this revertant cell clone might be similar to the 4R clone, a closer inspection of the peptide maps of the 866R clone revealed that a slight, but reproducible, difference did exist between the two revertant vole cell subclones. This difference reflected the presence of an additional peptide band immediately below the major C-terminal V8 protease fragment in the parental 866R revertant vole cell line (Figs 3 and 4b). Moreover, this particular peptide was absent in V8 protease hydrolysates of pp60^{src} from either chick cells infected with RSV rescued from the 866R cells or spontaneous re-transformants of this revertant vole cell clone (Fig. 4b). Although it may be involved in the morphological

Table 1 Quantitation of pp60^{src} protein and pp60^{src}-associated protein kinase activity in transformed and revertant vole cells

Cell line	pp60 ^{src} Protein (µg per mg cell protein)	Specific activity of pp60 ^{src} protein kinase (fmol ³² P incorporated per µg pp60 ^{src})	Total pp60 ^{src} protein kinase activity (fmol ³² P incorporated per mg cell protein)
Transformed 1T	0.30	135	40.5
Transformed 22T	0.07	383	26.8
Revertant 866R	0.21	321	64.2
Revertant 4R	0.13	259	33.7

Cultures were labelled with 25 µCi ml⁻¹ of ³⁵S-methionine for 2 h, and cell extracts were prepared as described previously^{3,7}. Samples were taken for the determination of protein content, and portions were immunoprecipitated with TBR serum. The amount of pp60^{src} protein in each extract was determined as previously described¹⁸. The resulting kinase activities, determined by quantitation of the phosphorylated IgG bands from polyacrylamide gels, were normalised with respect to the amount of cell extract protein used for the immunoprecipitation.

reversion of this cell line, the significance of this extra peptide is unclear. The presence of the peptide only in the revertant vole cell pp60^{src} and not in the *src* proteins derived from the rescued virus-infected (and transformed) chick cells and the re-transformed vole cells suggests that this alteration is not genetically stable and may be specific for the pp60^{src} of this revertant line. Nevertheless further studies are needed to determine whether this peptide represents expression of a *src* gene from a second ASV present in this particular revertant cell line.

pp60^{src}-associated protein kinase activity

As the pp60^{src} proteins from both revertant cell lines appeared structurally different from those obtained from the two transformed vole cell lines, we attempted to determine if they were functionally altered. It has been found⁶⁻⁸ that pp60^{src} is associated with a protein kinase activity. Using what is now considered to be the conventional assay for pp60^{src} protein kinase activity, we have found that both revertant, as well as both transformed, vole cell lines show pp60^{src}-protein kinase activity

(Table 1). No correlation can be made between the concentration of the *src* protein or the specific activity of pp60^{src}-protein kinase and the transformed or morphologically reverted state of the ASV-infected vole cells (Table 1).

Discussion

The slower electrophoretic mobility of the C-terminal V8 protease peptide of pp60^{src} from the revertant subclones represents a stable genetic change; however, it does not seem to impair the ability of the protein to induce morphological transformation. It is possible that the revertant subclones studied here were originally infected with a viral variant of SR-ASV or a different, contaminating strain of ASV, and that this is the reason for the differences in the pp60^{src} peptide maps. This seems unlikely as the immunoprecipitation with our TBR serum is strain specific for pp60^{src} of SR-ASV^{16,17}, and neutralisation data indicate that the rescued viruses obtained from the two revertant cell lines are the SR subgroup D strain of ASV. Peptide analyses on the *src* protein of other strains of ASV (Prague, Bratislava, Bryan) have shown that some variation in the V8 protease peptide maps does exist¹⁷. However, none of the viral pp60^{src} proteins studied so far have exhibited the peptide pattern (altered mobility of the C-terminal V8 protease peptide) of the proteins obtained from the two revertant vole cell lines studied here. It is of interest that a normal avian cell protein related to pp60^{src}, designated pp60^{src}, does have a C-terminal V8 protease peptide of the same electrophoretic mobility as those of the *src* proteins of the two revertant cell lines¹⁸. Additional subclones of revertants are now being isolated and peptide analyses of their pp60^{src} proteins being performed to determine if morphological reversion is always accompanied by the change in the C-terminal V8 protease peptide.

Although the pp60^{src} proteins in the transformed and revertant vole cells appear structurally different, they all exhibit protein kinase activity. It is possible that the *src* function involved in the changes in cellular morphology is not the pp60^{src}-associated protein kinase. It is also highly possible that pp60^{src} has other enzymatic activities. Alternatively the change in the *src* protein of the revertant cells may have altered *src* kinase specificity and/or cellular target recognition. Finally, the change in the pp60^{src} protein of the revertant cells may have nothing to do with the reversion process, but rather changes in the cells themselves have resulted in the ineffective action by the *src* protein in morphological transformation. For example, a cellular substrate of pp60^{src} may have been altered in such a way that it is no longer phosphorylated by the *src* kinase. This possibility appears highly plausible as the pp60^{src} protein from both revertant vole cell clones are able to cause morphological transformation in chicken cells and in retransformed vole cells. Thus, the transformed and revertant vole cells should provide an appropriate system to facilitate the identification of the target(s) of the ASV *src* gene product.

This investigation was supported by grants from the NIH (A.J.F., R.L.E.), American Cancer Society (R.L.E.) and Leukemia Research Foundation (A.J.F.). We also thank the Minnesota Leukemia Task Force for support. M.S.C. is a fellow in cancer research supported by a grant from the Damon Runyon-Walter Winchell Cancer Fund, and J.S.B. and A.F.L. are postdoctoral fellows supported by the NIH. R.A.K. is supported by a Young Investigator Award (CA 23833).

Received 9 April; accepted 30 July 1979.

- Hanafusa, H. in *Comprehensive Virology* Vol. 10 (eds Fraenkel-Conrat, H. & Wagner, R.) 401-483 (Plenum, New York, 1977).
- Vogt, P. K. in *Comprehensive Virology* Vol. 9 (eds Fraenkel-Conrat, H. & Wagner, R.) 341-455 (Plenum, New York, 1977).
- Brugge, J. S. & Erikson, R. L. *Nature* **269**, 346-348 (1977).
- Purchio, A. F., Erikson, E., Brugge, J. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1567-1571 (1978).
- Brugge, J. S., Erikson, E., Collett, M. S. & Erikson, R. L. *J. Virol.* **26**, 773-782 (1978).
- Levinson, A. D., Oppermann, H., Levintow, L., Varmus, H. E. & Bishop, J. M. *Cell* **15**, 561-572 (1978).
- Collett, M. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2021-2024 (1978).
- Erikson, E., Collett, M. S. & Erikson, R. L. *Nature* **274**, 919-921 (1978).
- Krzyzek, R. A., Lau, A. F., Faras, A. J. & Spector, D. H. *Nature* **269**, 175-179 (1977).
- Krzyzek, R. A., Lau, A. F., Vogt, P. K. & Faras, A. J. *J. Virol.* **25**, 518-526 (1978).
- Krzyzek, R. A., Lau, A. F. & Faras, A. J. *J. Virol.* **29**, 507-516 (1979).
- Lau, A. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Cleveland, D. W., Fischer, S. G., Kirschner, M. W. & Laemmli, U. K. *J. biol. Chem.* **252**, 1102-1106 (1977).
- Collett, M. S., Erikson, E. & Erikson, R. L. *J. Virol.* **29**, 770-781 (1979).
- Kessler, S. W. *J. Immun.* **115**, 1617-1624 (1975).
- Brugge, J. S., Steinbaugh, P. & Erikson, R. L. *Virology* **91**, 130-140 (1978).
- Brugge, J. S. *et al. J. Virol.* **19**, 1196-1203 (1979).
- Collett, M. S., Brugge, J. S. & Erikson, R. L. *Cell* **15**, 1363-1369 (1978).
- Boettiger, D. *Virology* **62**, 522-529 (1974).

letters

Is Z Cha another case of gravitational radiation?

H. Ritter

Max-Planck-Institut für Astrophysik, Föhringer Ring 6, D-8000 München 40, FRG

Z Cha is an eclipsing cataclysmic binary with an extremely short orbital period of only $P = 0.07449923$ d (ref. 1). Cataclysmic binaries are systems consisting of a white dwarf primary, of mass M_1 , and of a secondary of low mass M_2 , which fills its critical Roche volume². The secondary feeds an accretion disk around the primary by transferring matter through the inner lagrangian point L_1 . At the point where the streaming matter impacts the disk, a shock front is formed which is usually referred to as the hot spot. The geometry of the system, as seen from above the orbital plane, is shown in Fig. 1. The determination of the masses, the absolute dimensions and the mass exchange rate of Z Cha are reported here. Various possible mechanisms for the mass exchange are discussed which lead to the conclusion that it is probably due to the action of gravitational waves.

High-speed photometry observations¹ of Z Cha have clearly shown that two separate emitting regions undergo phase shifted total eclipses. These emitting regions are the hot spot and the centre of the accretion disk including the white dwarf primary. The half widths⁴ of the white dwarf eclipse, Δt_{WD} , and of the hot spot eclipse, Δt_{HS} —as well as their mutual phase shift δt —have been determined from improved photometric observations³. In addition the durations of the partial phases of the corresponding eclipses, Δ_{WD} and Δ_{HS} have been estimated⁵. The resulting numerical values of Δt_{WD} , Δt_{HS} , δt , Δ_{WD} and Δ_{HS} are listed in Table 1a. Recent spectroscopic observations⁶ have enabled the radial velocity K_1 and the projected orbital velocity of the accretion disk $v \cdot \sin i$ to be determined. These quantities are also listed in Table 1a.

The masses and the absolute dimensions of the system can be determined directly if the standard assumption⁷ is made, namely that $v \cdot \sin i$ is the keplerian rotation velocity at the hot spot's distance, d , from the centre of the disk. Analysis of the double eclipse using a method described elsewhere⁴ yields two equations for the mass ratio M_1/M_2 and the orbital inclination i in

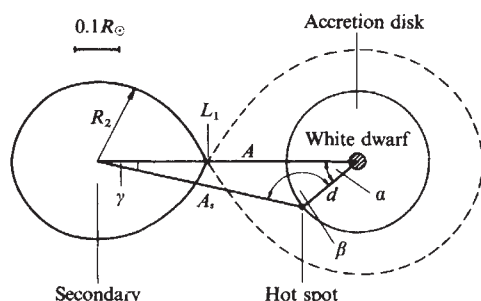


Fig. 1 Geometry of Z Cha as seen from above the orbital plane. Except for the radius of the hot spot the drawing is to scale.

terms of R_2/A and R_2/A_s . Using geometrical relations shown in Fig. 1 and those of the keplerian orbit these equations can be solved⁵. The resulting main parameters of Z Cha are listed in Table 1b. The radii of the white dwarf primary, R_1 , and of the hot spot, R_{HS} , have been computed from the durations of the partial phases of the corresponding eclipses⁵.

Most remarkably, although the secondary was not assumed to be a main sequence star (an assumption usually made when discussing cataclysmic binaries^{2,7}) it turns out that the secondary lies almost exactly on the theoretical mass–radius relation of low mass main sequence stars⁸. Even more surprisingly the primary fits the theoretical mass–radius relation of white dwarfs^{9,10}. Both these findings are not self-evident and must be considered as a confirmation of the reliability of the parameters given in Table 1. An additional consistency check is provided by comparing the value of α (see Fig. 1) given in Table 1b with the value $\alpha \approx 34^\circ$ derived from observations¹. The agreement is quite good. (Note that there apparently is an error in ref. 1 and that in fact $\alpha = 34^\circ$ and not $\alpha + \gamma = 34^\circ$; details will be given elsewhere⁵.) Moreover, if Δt_{HS} is replaced by α (34°) as an input parameter in the above analysis the values of all the resulting parameters are almost identical to those given in Table 1b.

It is now possible to estimate the rate of mass exchange, $\dot{M}$, by determining the luminosities of the hot spot and/or of the accretion disk. The rather lengthy procedure necessary to transform relative amplitudes of the observed light curves¹ into absolute luminosities is described elsewhere⁵. Basically the procedure involves four steps. First, the relative amplitudes of the different emitting regions, that is, of the secondary, the hot

Table 1 Parameters for Z Cha

a Observational data for Z Cha*			
Quantity	Symbol	Value	Ref.
Orbital period	P	0.07449923 d	1
	Δt_{WD}	(350 ± 10) s	3†
	Δt_{HS}	(585 ± 10) s	3†
	δt	(226 ± 10) s	3†
Defined in text	Δ_{WD}	(110 ± 10) s	3†
	Δ_{HS}	(190 ± 10) s	3†
	K_1	(122 ± 6) km s ⁻¹	6
	$2 \cdot v \cdot \sin i$	$(1,217 \pm 54)$ km s ⁻¹	6
b Resulting physical parameters for Z Cha			
Orbital inclination	i	$74^\circ \pm 0.8^\circ$	This paper
Mass ratio	M_1/M_2	2.2 ± 0.3	
Mass of the primary	M_1	$(0.35 \pm 0.06)M_\odot$	
Mass of the secondary	M_2	$(0.16 \pm 0.02)M_\odot$	
Radius of the primary	R_1	$(0.018 \pm 0.002)R_\odot$	
Radius of the secondary	R_2	$(0.19 \pm 0.01)R_\odot$	
Radius of the hot spot	R_{HS}	$(0.032 \pm 0.003)R_\odot$	
Orbital separation	A	$(0.59 \pm 0.03)R_\odot$	
Radius of the disk	d	$(0.16 \pm 0.01)R_\odot$	
	$\alpha \pm$	$39.8^\circ \pm 2.7^\circ$	
	$\beta \pm$	$127.5^\circ \pm 3.0^\circ$	

* Errors indicated are estimated 1σ (except for K_1 and $v \cdot \sin i$ (ref. 6)).

† Determined from the light curve given in ref. 3.

‡ Defined in Fig. 1.

spot and the accretion disk including the white dwarf primary, have been determined from observations of Z Cha at minimum light¹. Second, the measured amplitudes of the different emitting regions in 'white light' (ref. 1) have been converted to relative bolometric amplitudes, taking into account the different effective temperatures of the different emitting regions. Third, an absolute luminosity for one of the emitting regions must be determined. This has been done by assuming that the secondary is a normal main sequence star⁸ as indicated by its position in the mass-radius diagram. Finally the absolute luminosities of the hot spot and of the accretion disk have been computed and directly converted into accretion rates^{5,11}. The accretion rates determined in this way, $\dot{M}_{\text{HS}}$ from the hot spot luminosity and $\dot{M}_{\text{D}}$ from the disk luminosity, are listed in Table 2. The errors given are purely formal, the rates themselves are probably known only within a factor of 2.

Knowing the approximate accretion rate, what mechanism can give rise to a corresponding mass exchange rate? The most obvious mechanism, namely nuclear evolution of the secondary, can be ruled out because the secondary has a mass of only $\sim 0.16M_{\odot}$, and is almost certainly fully convective⁸ and thus remains chemically homogeneous throughout its evolution. This has the effect of decreasing the star's radius with time. Thus no mass exchange can be expected in these circumstances. However, even if the secondary were not fully convective it would increase its radius a nuclear time scale $\tau_{\text{nuclear}} \approx 10^{12}$ yr giving rise to a mass exchange rate of only $\dot{M} \approx 2 \times 10^{-13} M_{\odot} \text{ yr}^{-1}$. This is two orders of magnitude less than that estimated from observations.

The second likely mechanism, namely mass loss by the secondary on a thermal time scale, due to thermal readjustment, can also be ruled out. This is because the system, which has a mass ratio $M_1/M_2 \approx 2.2$, is secularly stable against mass exchange unless the mass-radius exponent of the secondary $\alpha_2 = d \log R_2 / d \log M_2$ is less than -0.7 . However, the corresponding value of α_2 for a $0.16M_{\odot}$ main sequence star is $\alpha_2 \approx +0.7$. In fact, even in the worst case of deviation from thermal equilibrium, that is when the fully convective star expands adiabatically, α_2 is only $-1/3$ and the system would still be secularly stable against mass exchange.

Thus the most obvious, classical mechanisms of mass exchange are insufficient to explain the observations of Z Cha. On the other hand, it has been suggested that cataclysmic variables, such as Z Cha, should exchange matter as a consequence of angular momentum loss due to gravitational radiation¹². (Note that this mechanism has the *a priori* advantage of always forcing the binary to some mass exchange independent of the mass-radius exponent of the mass-losing star.) According to Einstein's quadrupole formula¹³, gravitational waves should give rise to a mass exchange rate¹²

$$\dot{M}_{\text{GR}} = \frac{32G^3}{5C^5} (M_1 + M_2) M_1 M_2^2 A^{-4} \times \left[\frac{M_1 - M_2}{M_1} + \frac{1}{2} \left(\alpha_2 + \frac{M_1 + M_2}{M_1} \frac{d \log f_2}{d \log (M_1/M_2)} \right) \right]^{-1} \quad (1)$$

where f_2 is the Roche radius of the secondary in units of the

system's separation A . The numerical value of $\dot{M}_{\text{GR}}$ corresponding to the parameters listed in Table 1b is given in Table 2 and can thus be compared directly with the estimates $\dot{M}_{\text{D}}$ and $\dot{M}_{\text{HS}}$. Gravitational radiation apparently induces just that amount of mass exchange—within the uncertainties—estimated from observations. Moreover, unless one were prepared to accept rather exotic mechanisms for the mass exchange the only mechanism capable of explaining mass exchange at the observed level in a natural way is loss of angular momentum due to gravitational radiation. It can therefore be concluded that Z Cha is, in addition to the binary pulsar¹⁴, another example of the action of gravitational radiation.

I thank Drs F. Meyer and H.-C. Thomas for stimulating discussions and Dr J. J. Perry for improving the English.

Received 23 April; accepted 10 July 1979.

1. Warner, B. *Mon. Not. R. astr. Soc.* **168**, 235–247 (1974).
2. Robinson, E. L. *A. Rev. astr. Astrophys.* **14**, 119–142 (1976).
3. Fabian, A. C., Pringle, J. E., Whelan, J. A. J. & Bailey, J. A. *IAU Colloq.* No. 46 (in the press).
4. Ritter, H. & Schröder, R. *Astr. Astrophys.* **76**, 168–175 (1979).
5. Ritter, H. *Astr. Astrophys.* (in preparation).
6. Vogt, N. *Mitt. Astr. Ges.* (in the press).
7. Warner, B. *IAU Symp.* No. 73, 85–140 (1976).
8. Grossmann, A. S., Hays, D. & Graboske, H. C. Jr. *Astr. Astrophys.* **30**, 95–103 (1974).
9. Chandrasekhar, S. in *Stellar Structure*, 427 (University of Chicago Press, 1938).
10. Koester, D. *Astr. Astrophys.* **64**, 289–294 (1978).
11. Bath, G. T., Evans, W. D., Papaloizou, J. & Pringle, J. E. *Mon. Not. R. astr. Soc.* **169**, 447–470 (1974).
12. Faulkner, J. *Astrophys. J. Lett.* **170**, L99–L104 (1971).
13. Landau, L. D. & Lifshitz, E. M. *The Classical Theory of Fields*, 4th edn, 356 (Pergamon, Oxford, 1975).
14. Taylor, J. H., Fowler, L. A. & McCulloch, P. M. *Nature* **277**, 437–440 (1979).

Dissipation, merging and the rotation of galaxies

S. Michael Fall

Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge and Center for Astrophysics, Cambridge, Massachusetts

The distinction between elliptical and disk galaxies is often explained by invoking different rates of dissipation and star formation in the proto-galactic gas clouds from which they are presumed to have formed^{1–3}. Another possibility is raised by the suggestion that many ellipticals are the remnants of galaxies that merged while suffering the tidal distortion and violent relaxation of a slow encounter^{4,5}. Various *N*-body simulations indicate that merging can account for many of the observed properties of bright elliptical galaxies^{6–9}, although some reservations have been expressed about the possibility of accounting for their low rotation velocities¹⁰. My aim here is to consider a specific model for the formation of disk galaxies by dissipation and the formation of bright ellipticals by merging and to suggest that it can account for their rotation properties in a natural way.

The rotation of a self-gravitating body is conveniently expressed in terms of the dimensionless spin parameter $\lambda = J|E|^{1/2}G^{-1}M^{-5/2}$ where J , E and M are respectively the total angular momentum, energy and mass of the body and G is the gravitational constant. In the hierarchical clustering picture, all structures are endowed with some rotation before collapse as a consequence of their mutual tidal torques. The characteristic spin of this mechanism is $\lambda_T \approx 0.07$ and any significant differences from this value must be the result of non-gravitational effects^{11,12}. An elliptical galaxy has $\lambda \approx 0.3(v_r/\sigma_r)$ where v_r is the peak rotation velocity along the major axis and σ_r is the line-of-sight velocity dispersion near the centre of the galaxy⁶. The measurements of these velocities for 25 bright ellipticals^{13–17} give the average spin value $\lambda_E \approx 0.08$ and the distribution shown in Fig. 2 (neglecting small corrections for the

Table 2 Mass exchange rates

Quantity	Value ($\times 10^{-11} M_{\odot} \text{ yr}^{-1}$)
$\dot{M}_{\text{D}}^*$	1.9 ± 1.3
$\dot{M}_{\text{HS}}^\dagger$	4.3 ± 1.2
$\dot{M}_{\text{GR}}$	4.2 ± 1.6

* Estimated from luminosity of the accretion disk.

† Estimated from luminosity of the hot spot.

‡ Expected from the action of gravitational radiation.

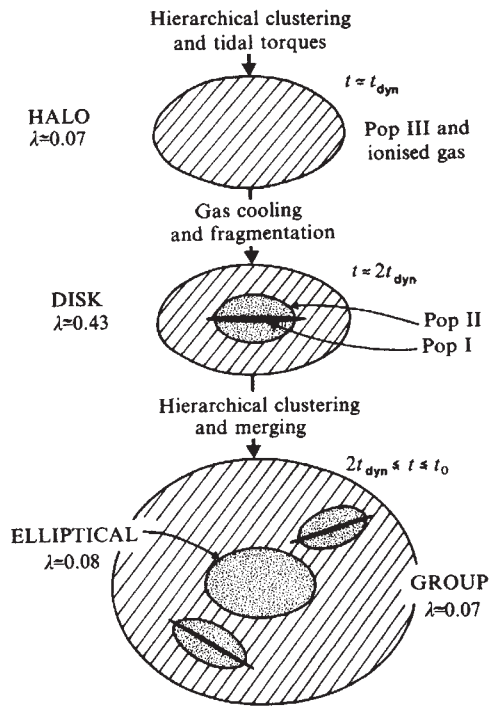


Fig. 1 Evolutionary sequence discussed in the text. Pre-galactic haloes are formed by the hierarchical clustering of dark stars and gas; disk galaxies result from dissipation and fragmentation within the haloes; elliptical galaxies are produced by merging in small groups.

orientations of the galaxies and uncertainties in the observations). Disk galaxies appear to be supported by rotation alone and, to the extent that this is true, have $\lambda_D \approx 0.43$. (This follows from equations 9, 11 and 14 of ref. 18, equation 6 on page 386 of ref. 19 and the virial theorem; it is exact for a cold exponential disk without a bulge component.)

The proposed evolutionary sequence, illustrated in Fig. 1, is based on the picture of galaxy formation and clustering outlined by White and Rees²⁰. In their view, pre-galactic haloes formed by gravitational clustering and consisted of low-mass or burnt-out stars (pop III) along with some residual gas. If the gas was originally ionised (for example by shocks), the condition that it could cool and collapse within the dark haloes sets their characteristic mass and (half-mass) radius at $M_H \approx 5 \times 10^{12} M_\odot$ and $R_H \approx 150$ kpc (to rough accuracy²⁰). The most likely outcome of this dissipative process would seem to be a population of disk galaxies²¹ (spirals if some gas was retained, lenticulars if it was depleted or removed). Their characteristic masses and radii must have been set by the poorly understood details of the collapse process and will therefore be parametrised in terms of the halo values: $M_D \approx x M_H$ and $R_D \approx y R_H$. On observational grounds^{18,22-24}, the values $x = y \approx 0.1$ are indicated and they satisfy the condition that the gas was able to reach a self-gravitating state before fragmenting into stars. The tidal spin of the haloes and the condition that angular momentum was conserved during the collapse of the gas now lead to the characteristic spin $\lambda_D \approx (xy)^{-1/2} \lambda_T$ for disk galaxies (neglecting geometrical factors of order unity). This estimate is based on the suggestion that any mechanisms for the transport of angular momentum (such as turbulent viscosity) would not have been efficient during collapse²⁵, and it agrees with the empirical spin value for $xy \approx 0.03$.

Gravitational clustering on scales larger than the individual haloes would not have been influenced by their internal dissipation and would have given rise to the observed hierarchy of groups and clusters^{26,27}. The characteristic masses and radii of the groups are related to the number of member galaxies m by the expressions $M_G \approx m M_H$, $R_G \approx m^{\alpha+1/3} R_H$ where $\alpha \approx 0.3-0.5$

is the index of the 'initial' perturbation spectrum $\delta\rho/\rho \propto M^{-\alpha}$. The groups of the first generation to reach equilibrium would have had about $m \approx 2-8$ members and tidal stripping is expected to have detached most of the haloes from their galaxies in a fraction $0.08 R_G/R_H \approx 0.08 m^{\alpha+1/3}$ of the first crossing time (equation A4 of ref. 20). This is expected to have left each group with a fairly smooth background of stripped halo material and to have left most of the luminous parts of the galaxies relatively undisturbed (so long as they are tightly bound). In this environment, the merging time scale for the galaxies themselves is about $0.03 x^{-1} (M_G/M_D) \approx 0.03 m x^{-2}$ group crossing time scales (equation A7 of ref. 20) so about one merger per group is expected before its incorporation into the next level of the hierarchy. After that, the conditions for merging are less propitious, especially in large clusters, because merging cannot occur if the encounter velocities of the interacting galaxies are more than a few times their internal velocity dispersions^{6,7} (~ 250 km s⁻¹). This condition also implies that the ellipticals produced from n disk galaxies have characteristic masses and radii with the approximate values $M_E \approx n M_D$ and $R_E \approx n R_D$.

The spins of elliptical remnants can include contributions from both the orbital and the internal angular momenta of the disk galaxies that produced them. To the extent that the merging process was hierarchical, the internal contribution is negligible because the galaxies most likely to merge were those that formed near each other and therefore provided the mutual torques responsible for their equal and opposite tidal spins. Some of the merging in large groups may, however, have involved galaxies that did not form near enough to guarantee this cancellation. If they had no correlation, the typical contribution was $n^{1/2} J_D$ to the angular momentum of a remnant and therefore $n^{-3/2} \lambda_D$ to its spin. Thus, even in this extreme case, the internal contribution becomes negligible after about two merger

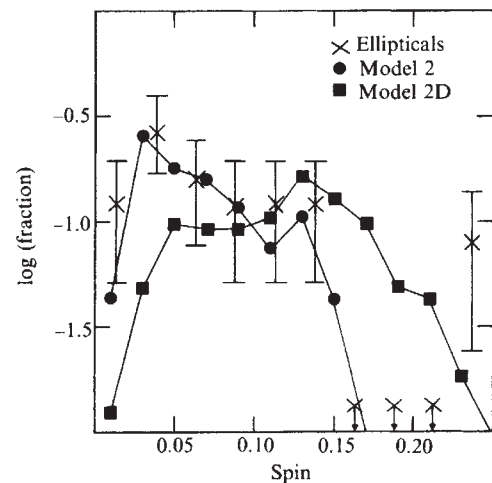


Fig. 2 Spin distributions. The normalised histogram for elliptical galaxies is from the data of refs 13-17 (listed also in Table 2 of ref. 9). The other histograms are for merger remnants at the end of two 1,000-body simulations that differ only in the initial spin values of their particles: $\lambda = 0.08$ for model 2 (one of the 'standard' simulations of ref. 9) and $\lambda = 0.40$ for model 2D (a new simulation with the same particle sizes, masses and orbits). The initial spin orientations are random and the final spin distributions are similar because of the statistical $n^{-3/2}$ dependence of the contribution from internal spins. Orienting the initial spins to cancel in pairs (by earlier tidal torques) would result in even more similar final distributions. The spin distributions are not sensitive to the sizes of the particles but they do depend somewhat on the initial position and velocity distributions (Poisson positions and Hubble velocities here). These simulations do not include a component of stripped halo material; however, the spins of merger remnants are not expected to depend sensitively on the presence of this component so long as it also participates in the hierarchical development of clustering.

events ($n \approx 3$). The orbital contribution to the spins of elliptical remnants must have derived from tidal torques in the clustering hierarchy, with a maximum value set by the condition that interacting galaxies should have overlapped enough to have merged within a few orbital periods. (Otherwise, the orbits could have been disrupted and merging prevented.) The simulations of merging pairs of galaxies^{6,7} indicate $\lambda_{\max} \approx 0.2$ (for parabolic encounters with zero net internal spin) and the cosmological simulations^{8,9} indicate that the average spin value of merger remnants is about $\frac{1}{2}\lambda_{\max} \approx 0.1$. This is because the distribution of impact parameters is fairly flat and most merging is from marginally bound orbits. Figure 2 shows the results of two simulations by Dr S. J. Aarseth and myself; the agreement with the spin distribution of real ellipticals is reasonable in both cases.

In many respects, the specific model outlined here is undoubtedly over-simplified and leaves some important questions unanswered: How did the dwarf ellipticals form? Why are ellipticals virtually gas-free? What determines the colours of elliptical remnants? The model suggests that different dissipation and star formation rates need not be invoked at the formation epoch of bright galaxies to explain their later morphological differences. The real situation may, however, have involved a variety of galaxy formation processes with merging only responsible for the making of some ellipticals. Perhaps the main conclusion to be drawn from the line of reasoning presented here is that the low rotation velocities of bright elliptical galaxies are no embarrassment to the idea that

they were produced by the merging of rapidly rotating disk galaxies.

I acknowledge helpful and interesting conversations with Sverre Aarseth, George Efstathiou, Martin Rees and Simon White.

Received 22 May; accepted 3 August 1979.

1. Sandage, A., Freeman, K. C. & Stokes, N. R. *Astrophys. J.* **160**, 831 (1970).
2. Gott, J. R. & Thuan, T. X. *Astrophys. J.* **204**, 649 (1976).
3. Binney, J. *Mon. Not. R. astr. Soc.* **177**, 19 (1976).
4. Toomre, A. & Toomre, J. *Astrophys. J.* **178**, 623 (1972).
5. Toomre, A. in *The Evolution of Galaxies and Stellar Populations* (eds Tinsley, B. M. & Larson, R. B.) 401 (Yale University Observatory, New Haven, 1977).
6. White, S. D. M. *Mon. Not. R. astr. Soc.* **184**, 185 (1978); *Mon. Not. R. astr. Soc.* (in the press).
7. Roos, N. & Norman, C. A. *Astr. Astrophys.* **76**, 75 (1979).
8. Jones, B. J. T. & Efstathiou, G. *Mon. Not. R. astr. Soc.* (in the press).
9. Aarseth, S. J. & Fall, S. M. *Astrophys. J.* (in the press).
10. White, S. D. M. *Astrophys. J. Lett.* **229**, L9 (1979).
11. Peebles, P. J. E. *Astrophys. J.* **155**, 393 (1969).
12. Efstathiou, G. & Jones, B. J. T. *Mon. Not. R. astr. Soc.* **186**, 133 (1979).
13. Bertola, F. & Capaccioli, M. *Astrophys. J.* **200**, 439 (1975).
14. Illingworth, G. *Astrophys. J. Lett.* **218**, L43 (1977).
15. Young, P., Sargent, W. L. W., Boksenberg, A., Lynds, C. R. & Hartwick, F. D. A. *Astrophys. J.* **222**, 450 (1978).
16. Davies, R. thesis, Cambridge Univ. (1978).
17. Schechter, P. L. & Gunn, J. E. *Astrophys. J.* **229**, 472 (1979).
18. Freeman, K. C. *Astrophys. J.* **160**, 811 (1970).
19. Watson, G. N. *A Treatise on the Theory of Bessel Functions* (Cambridge University Press, 1922).
20. White, S. D. M. & Rees, M. J. *Mon. Not. R. astr. Soc.* **183**, 341 (1978).
21. Larson, R. B. *Mon. Not. R. astr. Soc.* **176**, 31 (1976).
22. Eggen, O. J., Lynden-Bell, D. & Sandage, A. *Astrophys. J.* **136**, 748 (1962).
23. Ostriker, J. P., Peebles, P. J. E. & Yahil, A. *Astrophys. J. Lett.* **193**, L1 (1974).
24. Krumm, N. & Salpeter, E. E. *Astr. Astrophys.* **56**, 465 (1977).
25. Mestel, L. *Mon. Not. R. astr. Soc.* **126**, 553 (1963).
26. Peebles, P. J. E. *Astrophys. J. Lett.* **189**, L51 (1974).
27. Gott, J. R. & Rees, M. J. *Astr. Astrophys.* **45**, 365 (1975).

The vertical structure and thickness of Saturn's rings

Jeffrey N. Cuzzi*, Joseph A. Burns†, Richard H. Durisen‡ & Patrick M. Hamill§

* Ames Research Center, NASA, Moffett Field, California 94035

† Cornell University, Ithaca, New York 14850

‡ Indiana University, Bloomington, Indiana 47401

§ Langley Research Center, NASA, Hampton, Virginia

Observations of Saturn's rings over the last few years have contributed greatly to unravelling questions as to their nature and origin. However, some fundamental issues concerning the vertical structure and thickness of the rings have remained in dispute. In this note we present a synthesis of our results which reconciles most of the conflicting data. We first discuss the vertical structure and show how the many-particle-thick model appears to be feasible, but permits a local ring thickness of only a few tens of metres, much less than the thickness suggested by observations at ring-plane passage. We then describe how solar and satellite perturbations, which are coherent and not dispersive, do not significantly affect the true local vertical thickness of the rings, but do affect the tilt of the mean ring plane in a way that could be relevant to an understanding of the ring-passage observations.

One outstanding problem is the finding that the physical thickness of the saturnian rings is far too small to be directly observed. From measurements of the minimum brightness of the edge-on rings during ring-plane passage in 1966, ring thicknesses of 2.8 ± 1.5 km (ref. 1) and 1.4 ± 0.5 km (ref. 2) have been inferred. Reanalysis of these data³, however, gives a thickness of $0.8 (+2.3, -0.8)$ km, implying that only an upper limit on the thickness of ~ 3 km may be confidently claimed.

A further question concerns the local vertical structure of the

rings—are they many particles thick or are they essentially a monolayer? A ring that is many particles thick is the classical solution to the observed opposition effect⁴. Such a thickened ring implies random velocities with respect to the average local motion, and yet such velocities are continually damped by inelastic collisions between particles⁵. Such collisions occur about once per orbit for an optical depth of the order of unity. As a result, an originally 'thick' ring system flattens—in a time that is extremely short compared to 4.5×10^9 yr—to a layer that is at most a few particles thick unless the random motions are sustained by some ongoing source of energy. Obviously, in an essentially flat system, the ring thickness is of the order of the particle size. Thus, the above dynamical arguments and ring thickness 'observations' imply a monolayer with thickness (particle size) ~ 1 km; such a thickness is difficult to reconcile with either the opposition effect or with recent microwave observations^{6,7}.

Mechanisms considered as possible sources of energy for continual thickening of the rings include collisions among the particles^{4,8-10} and perturbations of the system by the saturnian satellites and the Sun^{4,10,11}. In recent studies^{12,13} we have discussed these and less significant mechanisms, such as meteoroid bombardment and radiation pressure, in detail.

Although, as argued above, collisions quickly flatten a ring system, the flattening proceeds, at most, only to the point where the system is a few particles thick. This is easily understood: imagine a ring system, nearly a monolayer, of identical particles of radius R in an orbit with semimajor axis a about a planet. Because of their keplerian motion, neighbouring particles will have a relative tangential velocity $\sim \omega R$, where ω = local orbital angular velocity. Particles collide because of this shear motion and their finite radii and soon develop an out-of-plane velocity component comparable to the shear velocity difference ωR over a particle diameter. By conservation of angular momentum in such collisions, the rings spread radially, in the absence of other restrictions, until the particles no longer collide. Such an evolution—rapid collisional flattening to a thickness of a few particle radii, followed by a very gradual spreading—was introduced by

Jeffreys⁵, treated numerically by Brahic⁹, and theoretically by Cook and Franklin⁷ and Goldreich and Tremaine¹⁰ with similar results. We also have investigated thickening mechanisms based on this 'Kepler shear' but applied to a ring system composed of particles of size distribution $n_0 R^{-3}$, which is that expected for an ensemble of particles which, perhaps only very early in their history, but possibly also at present, undergo comminution and mutual fragmentation. A size distribution of this general type seems to be required to explain the wavelength independence of the radar reflectivity of the rings^{6,7}. We have shown¹² that the ring system attains a steady state in which gravitational scattering by the large particles imparts random velocity components V_R to the small particles while inelastic particle collisions decrease these random velocities. This leads to a gaussian distribution $\exp(-z^2/z_0^2)$ of particle density with vertical displacement $z \sim a V_R/\omega a$.

Our results are not sensitive to the form of the size distribution. Suppose $n(R) = n_0 R^{-p}$ with minimum and maximum particle radii $R_{\min}$ and $R_{\max}$. The energy balance equation for the random particle motions is of the general form¹²

$$f_1 = \frac{\int M^2 n(M) dM}{\langle V_R \rangle^3}$$

where the constant f_1 includes the effects of inelastic collisions and Kepler shear, and depends only on ω , normal optical depth τ_0 , and coefficient of restitution e . We may obtain the particle mass distribution $n(M)$ from $n(R)$, and we obtain n_0 from τ_0 . Then,

$$n(M) = f_2 n_0 M^{-(p+2)/3}$$

where

$$n_0 = \frac{f_3(2-p)}{z_0(R_{\max}^{(3-p)} - R_{\min}^{(3-p)})}$$

and f_2, f_3 are constants.

Integrating over the size distribution, and using $z_0 = \sqrt{2} \langle V_R \rangle / \omega$, we obtain

$$z_0 = K R_{\max}, p \leq 3$$

$$z_0 = K \left(\frac{R_{\max}}{R_{\min}} \right)^{(3-p)/4} R_{\max}, 7 \geq p > 3$$

$$z_0 = K R_{\min}, p > 7$$

where K varies only slowly with p , τ_0 and e . The result is independent of the exponent p of the size distribution for $p \leq 3$. If $(R_{\max}/R_{\min}) \leq 10^2$ or 10^3 (ref. 6), z_0 is still comparable to $R_{\max}$ for $p \leq 4.5$. Evaluating the constant K for $p = 3$ (ref. 12), we find that $z_0 \sim 5 R_{\max}$.

Thus, although the full ring thickness is only about 10 times the radius of the largest particles, the ring is many small particles thick. Because the small particles dominate the surface area for an R^{-3} distribution, this can account for the optical observations^{4,14} and microwave scattering behaviour^{6,7}. An estimate of the maximum size of the ring particles following the R^{-3} size distribution can be obtained by showing that the gravitational energy given up in spreading the ring permits a ring thickness of only tens of metres to be sustained over the age of the Solar System. This implies that the maximum size particle in the power-law distribution is only a few metres in radius, similar to other estimates for uniform particle size cases^{9,10}. This true thickness is, however, much smaller than has been inferred^{1,2} from the ring-passage observations.

Dynamical perturbations of ring particles by the Sun and by Saturn's large satellites may be relevant to this discrepancy. Under the disturbing action of a single satellite along with the

planetary oblateness, the pole of a particle's mean orbit plane can be found to precess at a nearly uniform rate and at a nearly constant inclination relative to a mean pole, that of the Laplace plane¹³; for a ring particle, because it moves near the very oblate Saturn, the Laplace plane is almost, but not quite, coincident with the planet's equatorial plane. Following the usual celestial mechanics terminology, the displacement normal to the ring plane is decomposed into terms of different periodicities: short period ($T \sim$ particle orbit period), long period ($T \sim$ orbital precession period, orbital period of perturber) and secular ($T \sim$ infinity). The short period effects due to satellite and solar perturbations have an amplitude $\sim (m_j/m_s)(a/a_j)^3 \sin I_j$ (ref. 14) where m_j , a_j , and I_j refer to mass, semimajor axis, and inclination of the j th perturber. These amplitudes are of the order of 1 m and are coherent. Long-period and secular perturbations due to both the satellites and the oblateness of Saturn have been considered numerically⁴ and qualitatively¹⁶ by Bobrov. The principal effect of the oblateness is to slow or precess the particle orbit planes. However, Bobrov's¹⁶ amplitude of 300 m was taken from a numerical integration by Franklin¹¹, which did not include the planetary oblateness (F. A. Franklin, personal communication). Using a perturbation technique, we have obtained a closed form solution for the long-period and secular portion of the particle orbits' inclinations (measured relative to the Laplace plane) as perturbed by a satellite (or the Sun), including planetary oblateness and planetary precession¹³. Ignoring the latter effect, but including collisional damping, this forced inclination has constant amplitude $\omega_j \sin I_j (\omega_j + 2n_j)^{-1}$, and the forced precession of the orbit node is at the rate of $2n_j$, where n_j is the mean motion of the j th perturber, $\omega_j = Gm_j a_j^{-3/2} B_1^{3/2}(\alpha_j)/(4na^3)$ is the precession rate of the particle's orbit plane due to the j th satellite with G the gravitational constant, $\alpha_j = a/a_j$, $B_1^{3/2}(\alpha_j)$ is a Laplace coefficient¹³, and n is the particle's mean motion; $\omega_f = (3/2)J_2 n(r_0/a)^2$ is the precession rate of the particle's orbit plane due to the planetary oblateness represented by the gravitational coefficient J_2 as normalised to the saturnian radius r_0 . It is the inclusion of the effects of oblateness, manifested in ω_f , that effectively reduces the amplitude of the satellite perturbation compared with previous results^{4,11,16}. Typical amplitudes are less than a few metres due to the satellites, with Titan, Tethys, and Mimas being the only important ones, and ~ 10 m for the response to the solar perturbation which has the same character; these amplitudes increase by about an order of magnitude from the inner to the outer edge of the ring. A similar and still coherent effect is produced by the precession of Saturn's equatorial plane under solar torques, whereby particles in the mean ring plane experience noninertial effects causing particle orbit inclinations to oscillate while following the precession cycle of Saturn.

The subsequent evolution of such 'coherent' perturbations, in particular whether they may be coupled into random motions of the ring particles, has long been debated. The nature of the mathematical solution given by Burns *et al.*¹³ makes it fairly clear that there is no means by which this can occur. In contradiction to a previous suggestion¹⁶, differential precession of orbits of particles having similar inclination and slightly different semimajor axes, is not a property of the forced solution which dominates in the presence of damping (energy loss in collisions). Such 'winding up' does occur for a particular set of initial orbits randomly distributed about the mean ring plane (see equation (71) of ref. 13). 'Initial' noncircular components (random velocities) are, however, damped out on the flattening time scale, and the solution relaxes to the fully forced case which is coherent and time-independent (see for example equations (67) and (68) of ref. 13).

The above mechanisms produce displacements from the mean ring plane that are much less than the nominal 'observed' ring thickness. However, the Laplace plane mentioned above, about which the several metre coherent displacements occur, is itself significantly inclined to Saturn's equator by an amount that depends on the relative strength of the gravitational influence of Saturn and the various perturbing satellites, and hence on the

semimajor axis a of the particle orbit. Thus, the inclination of the mean ring plane varies with radial distance throughout the ring system, giving vertical displacements at the outer edges of the rings of about $\pm 10^2$ m each for perturbations due to Mimas, the Sun, and Titan, and smaller values for the other satellites. The nature of the result is a ring that is 'flipped' up and down at its opposite edges by possibly a total difference of several hundred metres, depending on the orbital parameters of the various perturbers. In addition, light scattered by the exterior, diaphanous E ring could also be misinterpreted as a 'thickness' of about 100 m or more¹³. Although the total effect is still much less than the 3 km 'upper limit' suggested by Lumme and Irvine³, it is not so far from their formal 'best result' of 800 m.

In conclusion, the 'vertical structure and thickness' of the rings encompass several distinct phenomena, produced by different causes and manifested in different observations. The quality that best describes the local microstructure of the rings is probably a true thickness of some tens of metres, caused by local, gravitationally induced, random particle motions; such a vertical structure is consistent with the opposition effect^{4,14} and with microwave observations^{6,7}. Satellite and solar perturbations produce two significantly different effects. First, coherent metre-scale vertical displacements, similar for all particles with similar orbital elements, are superposed on the local mean ring plane. There are both long-period and short-period displacements. The long-period displacements are characterised by particle orbits with a constant forced inclination and nodes that advance at twice the mean motion of the satellite perturber. The short-period displacements are tied to the longitude of the perturber. Second, the vertical displacement from the saturnian equatorial plane characterising the mean ring plane varies nonlinearly with radial distance and may produce an apparent edge-on thickness of several hundred metres, which would make it impossible for the true 'thickness' or local vertical scale height of the rings ever to be determined by ring-passage observations.

Our result—namely, that considerations of energy balance, combined with an R^{-3} power law with maximum radius $R_{\max}$, allow an approximate $R_{\max}$ of only several metres—has interesting implications for the origin of the rings. Particles that condensed directly from a protoplanetary nebula might be expected to be of this size, whereas residual particles resulting from Roche breakup of a large object would be expected to be significantly larger^{16,17}.

We acknowledge helpful conversations with A. Brahic, F. Franklin, P. Goldreich, R. Greenberg, R. Miller, and J. Pollack. Discussions with W. Ward concerning interpretation of our analytical results were particularly helpful. This work was initiated when the authors were at Ames Research Center, NASA; J.A.B. on an Intergovernmental Personnel Exchange Act, and J.N.C. and R.H.D. as NAS-NRC Resident Research Associates; P.H. was supported partly under NASA grant NSG-2227.

Received 19 April; accepted 6 August 1979.

1. Focas, J. H. & Dollfus, A. *Astr. Astrophys.* **2**, 251 (1969).
2. Kiladze, R. I. *Byull. Abastumani Astrofiz. Obs.* **37**, 151 (1969).
3. Lumme, K. & Irvine, W. M. *Astr. Astrophys.* **71**, 123 (1979).
4. Bobrov, M. S. in *Surfaces and Interiors of Planets and Satellites* (ed. Dollfus, A.) 377 (Academic, New York, 1970).
5. Jeffreys, H. *Mon. Not. R. Astr. Soc.* **107**, 263 (1947).
6. Pollack, J. B. *Space Sci. Rev.* **18**, 3 (1975).
7. Cuzzi, J. N. & Pollack, J. B. *Icarus* **33**, 233 (1978).
8. Cook, A. F. & Franklin, F. A. *Astr. J.* **69**, 173 (1964).
9. Brahic, A. *Astr. Astrophys.* **54**, 895 (1977).
10. Goldreich, P. & Tremaine, S. *Icarus* **34**, 227 (1978).
11. Franklin, F. A. in *The Rings of Saturn* (eds Palluconi, F. & Pettengill, G. H.) 178 (NASA SP-343, 1974).
12. Cuzzi, J. N., Durisen, R. H., Burns, J. A. & Hamill, P. *Icarus* **38**, 54–68 (1979).
13. Burns, J. A., Hamill, P., Cuzzi, J. N. & Durisen, R. H. *Astr. J.* (in the press).
14. Kawata, Y. & Irvine, W. M. in *Exploration of the Planetary System* (eds Woszczyk, A. & Iwaniszewska, C.) 441, (1974).
15. Vaaranen, P. *NASA Publ. N73-19854; Report 14, Aarne Karjalainen Obs.* (Univ. Oulu, Oulu, Finland, 1973).
16. Bobrov, M. S., *Nature* **273**, 284 (1978).
17. Harris, A. W. *Icarus* **24**, 190 (1975).
18. Greenberg, R., Davis, D. R., Hartmann, W. K. & Chapman, C. R. *Icarus* **30**, 769 (1976).

Kinetics of electron attachment to oxygen and water in flames

J. M. Goodings* & A. N. Hayhurst

Department of Chemical Engineering and Fuel Technology, University of Sheffield, Mappin Street, Sheffield, UK

Negative ions are found in flames¹ and are presumed to arise from electrons attaching to some unidentified flame species, possibly molecular oxygen^{2,3}. Free electrons and positive ions are produced in a flame either by chemi-ionisation⁴ in $\text{CH} + \text{O} \rightarrow \text{CHO}^+ + e^-$ or by collisional ionisation⁵ of a metal, such as potassium, in $\text{K} + \text{M} \rightarrow \text{K}^+ + e^- + \text{M}$, where M represents a third body. An atmospheric pressure flame is an interesting medium in which to measure the rate of appearance of negative ions; normally no more than one particle in 10^6 is electrically charged, there are usually no electric or magnetic fields and the attaching electrons are 'thermalised', that is, their energies are governed by the flame temperature. In addition, electron attachment is at much higher temperatures and pressures than normal^{6,7} and in a system free from wall effects. We present here a study of the kinetics of electron attachment to oxygen and water in flames. The rate coefficients suggest a surprisingly large negative temperature coefficient.

Most of this work was done with an oxygen-rich atmospheric-pressure flame of unburnt composition: $\text{H}_2/\text{O}_2/\text{N}_2 = 1.5:1.0:4.67$ and temperature 1,810 K. The burner⁸ gave a cylindrical flame (diameter 12.2 mm) with an almost flat reaction zone, followed by a burnt gas region in which there was laminar and plug flow, with conditions on the axis being unaffected by the surrounding atmosphere for around the first 25 mm. The mole fractions of the major constituents of the burnt gases were: $\text{O}_2 = 0.0379$; $\text{H}_2\text{O} = 0.233$; $\text{N}_2 = 0.726$. In addition, radicals were present with mole fractions for equilibrium in the burnt gases of: $\text{H} = 2.97 \times 10^{-6}$; $\text{OH} = 6.80 \times 10^{-4}$; $\text{O} = 2.42 \times 10^{-5}$; $\text{HO}_2 = 2.45 \times 10^{-7}$. Concentrations of H atoms were measured on a slightly different burner by an adaptation⁹ of the Li/Na comparison method¹⁰. The results were expressed as a disequilibrium parameter $\gamma_{\text{H}} (= [\text{H}]/[\text{H}]_{\text{e}})$, that is, the extent by which hydrogen atoms exceed their equilibrium concentration $[\text{H}]_{\text{e}}$ for the almost constant temperature of the burnt gases. It can be shown that $\gamma_{\text{H}}^{1/3} = \gamma_{\text{OH}} = \gamma_{\text{O}}^{1/2}$. Values of γ_{OH} in this flame were ~ 8 close to the reaction zone and ~ 1.4 at 20 mm downstream, after a time elapse of 3 ms. Because the axial velocity (7.0 m s^{-1}) in the burnt gases is well defined, observations of ion concentrations at different distances enable kinetic measurements to be made.

Free electrons were produced by atomising an aqueous solution of KOH into the burner supplies. Ion concentrations were measured, as previously described^{11,12}, by continuous sampling from points along the burnt gases into a quadrupole mass spectrometer. The major ions were K^+ , $\text{K}^+\text{H}_2\text{O}$, OH^- , $\text{OH}^-\text{H}_2\text{O}$ and O_2^- . Their abundances along the flame are given in Fig. 1, where the concentrations of hydrates have been included in those of the parent ions, because hydrates are largely formed during sampling¹³. Charge neutrality enables the electron concentration to be derived from $[e^-] = [\text{K}^+] - [\text{Y}^-]$, with $[\text{Y}^-] = [\text{OH}^-] + [\text{O}_2^-]$, that is, the total concentration of negative ions. Figure 1 shows that $[e^-]$ and $[\text{K}^+]$ rise much more rapidly than $[\text{OH}^-]$. This is consistent with K atoms undergoing collisional ionisation to give K^+ and free electrons, followed by electron attachment yielding O_2^- and OH^- . The ratio $[\text{OH}^-]/[\text{O}_2^-]$ increases with distance to a steady level above that calculated for equilibrium. That the ratio initially rises suggests

* Present address: Department of Chemistry, York University, 4700 Keele Street, Downsview, Ontario, Canada M3J 1P3.

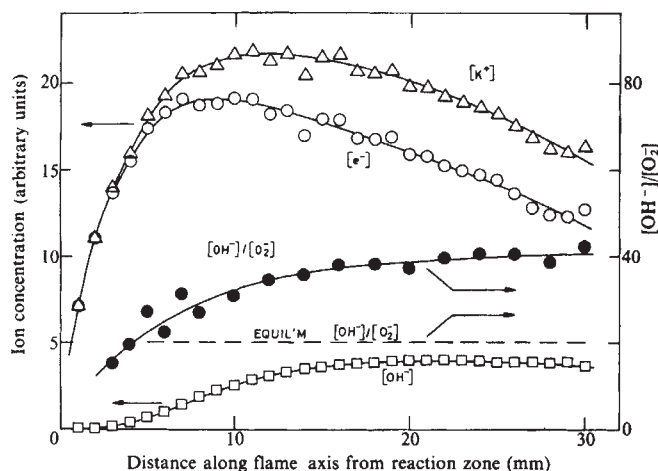
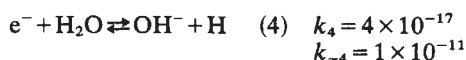


Fig. 1 Ion concentrations along the axis of the flame described in the text with 3×10^{-3} M KOH being sprayed in the atomiser. Also shown is the observed $[\text{OH}^-]/[\text{O}_2^-]$, together with its equilibrium value of 20.1 (dashed line).

that OH^- is produced from O_2^- . The near constancy of the ratio indicates that some rapid reaction couples the concentrations of OH^- and O_2^- , coming close to a local balance or equilibrium after a few millimetres. In contrast, the ratios of either $[\text{OH}^-]$ or $[\text{O}_2^-]$ to $[\text{e}^-]$ are not constant, being smaller than their equilibrium values, which in fact are approached well downstream. All this indicates that electron attachment giving O_2^- and OH^- is slow enough here to be a kinetic, rather than a steady-state, situation.

The possible formation and loss reactions for OH^- and O_2^- are:



where k_x and k_{-x} are the rate constants of the forward and reverse steps of reaction (x), whose equilibrium constant $K_x = k_x/k_{-x}$. The quoted values are those recommended by Jensen and Jones¹⁴ for 1,810 K in the $\text{ml molecule}^{-1} \text{s}^{-1}$ system of units. Reaction (3) can be discounted in our flame, as not being competitive with (1); similarly, step (−3) is much slower than (−1) because the electron affinity of OH exceeds that of O_2 . It might be noted that reaction (3) is the sum of (1) and (2), so that $K_3 = K_1K_2$. The value given for k_{-5} refers to room temperature¹⁵; k_{-6} has not been measured, although steric effects might be expected to make it smaller than k_{-4} and k_{-5} , for which the neutral reagent is atomic. Reaction (2) is the most likely inter-conversion process for OH^- and O_2^- , in which case Fig. 1 indicates it is close to being balanced a few mm from the reaction zone. Also, the initial rise in $[\text{OH}^-]/[\text{O}_2^-]$ is consistent with O_2^- being the first negative ion formed; that this ratio exceeds its equilibrium value is because $\gamma_{\text{OH}} > 1$ and, more importantly, reaction (2) is fast enough to proceed in the exothermic direction during sampling⁸. This latter conclusion is confirmed by the fact that the observed $[\text{OH}^-]/[\text{O}_2^-]$ depended slightly on the diameter of the sampling orifice. This compares with the ratio $[\text{e}^-]/[\text{Y}^-]$ being independent of hole size, in line with a relatively slow production of negative ions Y^- by electron attachment.

As a first step, consider only reactions (1) and (2), so that:

$$d[\text{Y}^-]/dt = k_1[\text{e}^-][\text{O}_2][\text{M}] - k_{-1}[\text{O}_2^-][\text{M}] \quad (\text{I})$$

This may be expressed as:

$$y = \frac{v}{[\text{e}^-]} \frac{d[\text{Y}^-]}{dz} = k_1[\text{O}_2][\text{M}] - \gamma_{\text{OH}}^{-1} \frac{[\text{OH}^-]}{[\text{e}^-]} \frac{k_1[\text{O}_2][\text{M}]}{K_3[\text{OH}]_e} \quad (\text{II})$$

with the flame velocity $v = dz/dt$ and $[\text{O}_2^-] = [\text{OH}^-][\text{O}_2]/(K_2[\text{OH}]_e\gamma_{\text{OH}})$, provided (2) is balanced. All this is checked in Fig. 2a, which is a plot of y against $x_1 = \gamma_{\text{OH}}^{-1}[\text{OH}^-]/[\text{e}^-]$. For this, values of y were obtained by numerically smoothing and differentiating¹⁶ the distance profile of $[\text{Y}^-]$. It is seen in Fig. 2a that y initially rises and is then linear in x_1 from 9 to 16 mm from the reaction zone. The central linear portion yields $k_1 = 3.2 \times 10^{-34}$ from its y intercept and the different $k_1 = 4.5 \times 10^{-34} \text{ ml}^2 \text{ molecule}^{-2} \text{s}^{-1}$ from the slope.

The departures from linearity are understandable in terms of minor contributions from reactions (4)–(6). The inclusion of reaction (4), along with (1) and (2), gives the rate equation:

$$y = k_1[\text{O}_2][\text{M}] + k_4[\text{H}_2\text{O}] - \frac{[\text{OH}^-]}{[\text{e}^-]} \times \\ \times (\gamma_{\text{OH}}^{-1}k_1[\text{O}_2][\text{M}] + \gamma_{\text{OH}}^3k_4[\text{H}_2\text{O}]) \frac{1}{K_3[\text{OH}]_e} \quad (\text{III})$$

where $[\text{H}]_e/K_4 = [\text{H}_2\text{O}]/(K_3[\text{OH}]_e)$. Letting $r = k_4[\text{H}_2\text{O}]/k_1[\text{O}_2][\text{M}]$ be the ratio of the forward rates of reactions (4) and (1), an estimate of r can be obtained from Fig. 2a where $y = 0$. This is at $z = 22$ mm from the reaction zone, where $\gamma_{\text{OH}} = 1.32$ and $[\text{OH}^-]/[\text{e}^-]$ has approximately its equilibrium value of $K_3[\text{OH}]_e = 0.250$. (This latter observation confirms that our measurements of $[\text{OH}^-]$ and $[\text{e}^-]$ are free from sampling falsifications^{8,13}.) Thus, equation (III) may be solved for the derived value of $r = 0.185$, whose magnitude shows electron attachment in (4) to be much less than via (1). In any case, equation (III) becomes:

$$y = (1+r)k_1[\text{O}_2][\text{M}] - (\gamma_{\text{OH}}^{-1} + r\gamma_{\text{OH}}^3) \frac{[\text{OH}^-]}{[\text{e}^-]} \frac{k_1[\text{O}_2][\text{M}]}{K_3[\text{OH}]_e} \quad (\text{IV})$$

Figure 2b shows a plot of y against $x_2 = (\gamma_{\text{OH}}^{-1} + r\gamma_{\text{OH}}^3)([\text{OH}^-]/[\text{e}^-])$ with $r = 0.185$. There is now a more extensive region of linearity from 9 to 24 mm, which yields $k_1 = 6.6 \times 10^{-34}$ from the y intercept and $k_1 = 6.4 \times 10^{-34} \text{ ml}^2 \text{ molecule}^{-2} \text{s}^{-1}$ from the slope, that is, differing by only 3%. The mean value of $k_1 = 6.5 \times 10^{-34}$ together with $r = 0.185$ gives $k_4 = 7.9 \times 10^{-17} \text{ ml molecule}^{-1} \text{s}^{-1}$, which is in excellent agreement with Jensen and Jones's estimate¹⁴, based largely on a room temperature measurement¹⁷ of k_{-4} . The occurrence of points in Fig. 2b off the straight line is thought to reflect slight inaccuracies in γ_{OH} , which appears as $\gamma_{\text{OH}}^3 (= \gamma_{\text{H}})$ in equations (III) and (IV), rather than deficiencies in the kinetic model. For example, the point

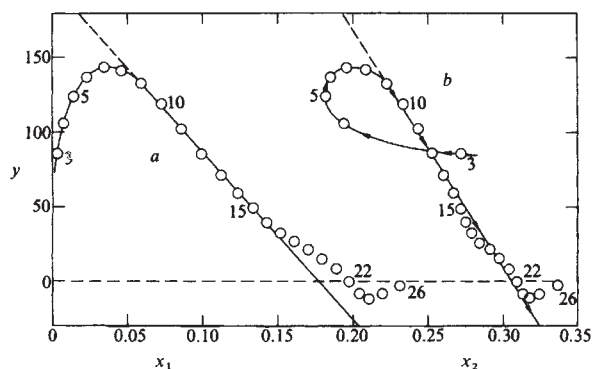


Fig. 2 The data in Fig. 1 plotted as $y = v/[\text{e}^-] \times d[\text{Y}^-]/dz$ against $x_1 = \gamma_{\text{OH}}^{-1}[\text{OH}^-]/[\text{e}^-]$ in a and against $x_2 = (\gamma_{\text{OH}}^{-1} + r\gamma_{\text{OH}}^3)[\text{OH}^-]/[\text{e}^-]$ in b. Points are numbered according to their distance from the reaction zone in mm and the arrows in b denote the sequence of points at increasing times in the burnt gases.

5 mm from the reaction zone becomes on the straight line, if γ_{OH} there is changed from 2.84 to 3.08, which is just within the error bounds for γ_{OH} . Figure 2b shows that γ is initially small, then increases and finally becomes negative. That $[Y^-]$ first shifts away from and then back through equilibrium is mainly caused by γ_{OH} declining, and consequently electron detachment being initially dominated by (-4) , then becoming less important and finally increasingly by (-1) .

A detailed examination shows that the effect of including process (5) is to shift those experimental points for conditions near the reaction zone farther away from the straight line in Fig. 2b. We thus conclude that the contribution from reaction (5) is insignificant. Moreover, similar reasoning applied to reaction (6) indicates negligible participation, particularly because $[H_2O_2]_e$ is extremely small at flame temperatures. Finally, the derived k_1 and k_4 were found not to depend systematically on either the diameter of the sampling orifice or on the amount of potassium added to the flame. This establishes that our measurements are free from any falsifying perturbations encountered^{8,18} with fast reactions proceeding during sampling.

In conclusion, we find $k_4 = 8 \pm 4 \times 10^{-17}$ ml molecule⁻¹ s⁻¹, which is very much in line with estimates based on room temperature data^{14,17}. Also, the behaviour of $[OH^-]/[O_2]$ in Fig. 1 is consistent with $k_2 > 10^{-12}$ ml molecule⁻¹ s⁻¹, which is not surprising. On the other hand, our conclusion that $k_1 = 6 \pm 3 \times 10^{-34}$ ml² molecule⁻² s⁻¹ is much less than the lower

temperature measurements of $k_1 \sim 10^{-31} - 10^{-30}$ for a reaction not expected to be very temperature-dependent^{6,7,14}. That reaction (3) is not participating indicates that k_3 must also be much smaller than the estimated¹⁴ value of 3×10^{-31} at $\sim 2,000$ K. We can only conclude that three-body electron attachment in a process like (1) or (3) has a surprisingly large negative temperature coefficient.

We thank J. E. Ogden for help with measurements of flame radical concentrations and Dr N. A. Burdett for helpful discussions.

Received 25 June; accepted 30 July 1979.

1. Gaydon, A. G. & Wolfhard, H. G. *Flames, their Structure, Radiation and Temperature*, 3rd edn (Chapman and Hall, London, 1970).
2. Goodings, J. M., Bohme, D. K. & Ng, Chun-Wai *Combust. Flame* (in the press).
3. Hayhurst, A. N. & Kittelson, D. B. *Combust. Flame* **31**, 37 (1978).
4. Green, J. A. & Sugden, T. M. *9th. Int. Symp. Combust.* 607 (1963).
5. Ashton, A. F. & Hayhurst, A. N. *Combust. Flame* **21**, 69 (1973).
6. Pack, J. L. & Phelps, A. V. *J. chem. Phys.* **44**, 1870 (1966); **45**, 4316 (1966).
7. Phelps, A. V. *Can. J. Chem.* **47**, 1783 (1969).
8. Burdett, N. A. & Hayhurst, A. N. *Proc. R. Soc. A* **355**, 377 (1977).
9. Goodings, J. M., Hayhurst, A. N. & Ogden, J. E. (in preparation).
10. Bulewicz, E. M., James, C. G. & Sugden, T. M. *Proc. R. Soc. A* **235**, 89 (1956).
11. Hayhurst, A. N., Mitchell, F. R. G. & Telford, N. R. *Int. J. Mass Spec. Ion Phys.* **7**, 177 (1971).
12. Hayhurst, A. N. & Telford, N. R. *Combust. Flame* **28**, 67 (1977).
13. Hayhurst, A. N. & Telford, N. R. *Proc. R. Soc. A* **322**, 483 (1971).
14. Jensen, D. E. & Jones, G. A. *Combust. Flame* **32**, 1 (1978).
15. Fehsenfeld, F. C., Ferguson, E. E. & Schmeltekopf, A. L. *J. chem. Phys.* **45**, 1844 (1966).
16. Hershey, H. C., Zakin, J. L. & Simha, R. *Ind. Eng. Chem. (Fund.)* **6**, 413 (1967).
17. Ferguson, E. E., Fehsenfeld, F. C. & Schmeltekopf, A. L. *Adv. Chem. Ser.* **80**, 83 (1969).
18. Hayhurst, A. N. & Kittelson, D. B. *Combust. Flame* **28**, 137 (1977).

Combined rotating and linear motion effects on the wear of polymers

B. J. Briscoe & T. A. Stolarski

Department of Chemical Engineering and Chemical Technology,
Imperial College, Prince Consort Road, London SW7, UK

The unusually low friction of polytetrafluoroethylene, PTFE, is utilised in many commercial applications¹. Unfortunately this low friction is often accompanied by rather high wear, and it is generally necessary to modify the polymer, usually by the incorporation of solid inorganic filler particles, to reduce the wear to acceptable levels^{1,2}. Although several attempts have been made to establish the mechanism of this improved wear behaviour, the optimisation of filler type, size and concentration remains an empirical exercise, and several relatively simple wear testing machines have been developed to assess prospective materials^{3,4}. These tests attempt to reproduce the sliding conditions encountered in practice. A widely used test is the pin-on-disk configuration where a fixed pin of polymer is loaded against a rotating disk. The load and relative velocity are fixed and the pin describes a circular path over the disk surface for a pre-determined period, after which the wear is measured, generally by weighing the polymer pin. Important additional variables are the environment, the surface topography of the disk, the temperature and the geometry of the pin⁵. In the process of conducting such a series of measurements we have identified and report here an additional variable which may be of general importance: the radius, R , of the path described by the pin on the disk.

Figure 1 shows data for the wear per unit distance, K , for specimens of PTFE and low density polyethylene, LDPE, as a function of R obtained using a pin-on-disk machine (for experimental details see ref. 6). We observe that K is a marked function of R for PTFE but not for LDPE. Figure 2 shows the same data and also data for other polymers as well as carbon-filled PTFE. Here, the wear data have been normalised by dividing the $K(R)$ obtained at a particular value of R by the value for $K(\infty)$ which is obtained by extrapolating $K(R)$ against $1/R$ curves to $1/R = 0$. The value of $K(\infty)$ corresponds to the

wear rate for 'pure' linear sliding. In Fig. 2 the data are presented as the ratio of $K(R)/K(\infty)$ against $1/R$ which has the virtue of superimposing the sets of data for PTFE and linear polythene on two master curves. The sliding velocities, V , and nominal contact pressures are constant for each specimen. The disk is cooled by a coolant flow beneath the disk surface and the bulk temperature is $\sim 20^\circ\text{C}$ (ref. 6). Note that two polymers, polymethylmethacrylate, PMMA, and LDPE, show no radius effect until R approaches the radius of the pin. However, for PTFE and a carbon-filled PTFE composite the wear decreases significantly as R is decreased. The data for high density polythene, HDPE, and ultrahigh molecular weight polythene, UHMPE, show a similar but less marked effect. Simultaneous measurements of the frictional force have also been made, and these data indicate that where the wear decreases the friction increases.

The common reduction in wear as R approaches the pin radius is due to the fact that, although the polymer surface is damaged, the comparatively small relative displacement of the surfaces enables much of the wear debris to be confined in the contact zone. The weight loss from the pin is thus reduced. At larger values of R the unusual behaviour of PTFE and, to a lesser extent, the linear polythenes seem to arise from a peculiarity in the pin-on-disk method. Since the pin describes a circular path the motion is complex and may be considered to be comprised of a linear vector, V , combined with a pin rotating with an effective angular velocity, ω_p , equal to that of the disk, ω_d . As R decreases while V is maintained constant, ω_p increases as $1/R$.

It is this angular component which we believe is responsible for the unusual friction and wear behaviour shown by PTFE and the linear polythenes. The reason for this effect is not clear, but the following rationale accounts for the results. PTFE and HDPE have unusual friction and transfer behaviour although the behaviour is more pronounced by PTFE^{7,8}. If a PTFE pin is slid linearly at low speeds against a clean smooth surface, the static frictional force is initially high and a relatively thick transfer layer is deposited. Further sliding produces a thinner and more highly orientated film and a lower dynamic frictional force. If sliding is arrested and the polymer pin is then rotated about its load axis by 90° , the initial behaviour is repeated⁷. We have also conducted a series of measurements where polymer spheres of ~ 8 mm diameter have been subjected to a simul-

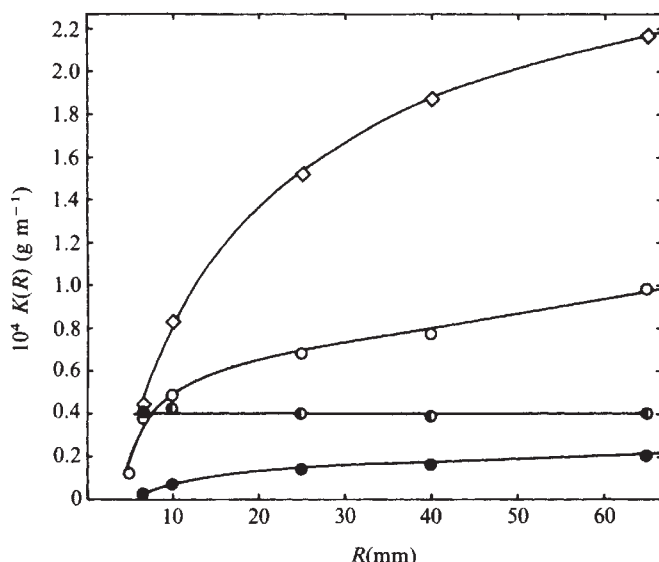


Fig. 1 Rate of wear, $K(R)$, expressed as mass loss per unit sliding distance, as a function of R , the radius of the wear path for PTFE and LDPE. Sliding velocity: 0.5 m s^{-1} ; nominal contact pressure: $6 \times 10^5 \text{ Pa}$; counterface temperature $\sim 20^\circ \text{C}$; counterface mild steel, roughness $\sim 0.1 \mu\text{m c.l.a.}$; pin diameters: PTFE $\bullet$, 6 mm; $\circ$, 13 mm; $\diamond$, 20 mm; LDPE $\bullet$, 13 mm (wear rate $\times 10^2$). The wear rate is a marked function of R .

taneous combination of linear and rotating motion in an attempt to reproduce the effective motion of the pin in the wear testing machine. The frictional force in the direction of linear motion was found to be extremely sensitive to the introduction of rotation for PTFE materials and also to a lesser extent for HDPE and UHMPE. The frictional force in this direction typically increases from the low kinetic value to a maximum value close to the static value as the imposed rotational velocity is increased, and it reaches a maximum when the angular velocity is such that the contact area is rotated by $\pi/2$ for a linear travel approximately equal to the contact diameter. This result is consistent with the response when a linear motion is interrupted and an instantaneous $\pi/2$ motion is imposed on the contact. When repeated traversals of the fixed polymer pin are made over the transferred layer, the pin seems to displace the layer readily and a further layer is deposited. This process leads to high wear and there is evidence that the rate of wear is greatly influenced by the magnitude of the adhesion of the transferred layer to the substrate^{8,9}; the greater the strength of adhesion the lower the rate of wear. When rotation is introduced during linear sliding the transferred films seem to be somewhat thicker and more tenaciously attached to the counterface for PTFE, although we have no quantitative results. These effects are restricted to sliding on smooth counterfaces and apparently to contacts where either the sliding velocity or the contact pressures have moderate values. For a large apparent contact pressure of $\sim 1.3 \times 10^7 \text{ Pa}$ the unusual response decays quickly at sliding velocities in $>20 \text{ mm s}^{-1}$. At a more moderate contact pressure of $6 \times 10^5 \text{ Pa}$, which is typical of commercial contacts, the unusual behaviour is maintained to above 2.0 m s^{-1} . When the counterface is rough, for example when the centre line average roughness is $\sim 3.0 \mu\text{m}$, the unusual friction behaviour of PTFE is absent in all the contact conditions studied.

In contrast LDPE shows only high friction and thick film behaviour with no orientation effects. This polymer also shows no increase in the frictional force as measured in the linear sliding direction when rotation is imposed. PMMA is similarly insensitive. The special friction and wear behaviour seems to be restricted to 'smooth molecular profile' polymers when they are slid on smooth counterfaces and in no more than moderately intense contact conditions. These effects seem to be associated with reorientation processes occurring in the contact. The

introduction of rotation into a linear sliding motion will clearly affect this process and hence produce thicker transferred films. The friction data observed in the pin-on-disk experiment are consistent with this, as are the low speed data. To explain the reduction in wear with small additions of angular velocity we must then assume that the thick unorientated layers adhere more strongly to the substrate than the thinner more highly orientated films⁷ and hence a lower wear results. The comparatively low wear of LDPE compared with PTFE and HDPE in comparable conditions is consistent with this hypothesis.

The results on the pin-on-disk machine indicate that the radius effect described may be interpreted if we assume that the polymer pin is effectively rotated as it describes its circular path of radius R . We have conducted a series of measurements where the pin is actually rotated about its load axis while it is in contact with the rotating disk. The direction and magnitude of this pin rotation, ω_p , have been varied over a wide range for a fixed value of R . When the value of ω_p is equal in magnitude to ω_d but rotating in the opposite direction the wear rate of PTFE is observed to have its maximum value. This corresponds to effectively 'pure' linear sliding. If, however, the pin is rotated in the same direction as that effectively induced by the disk's rotation a minimum in the wear rate is observed when ω_p is such that the total pin rotation is equal to a value of $\pi/2$ for a linear travel distance of approximately one contact diameter. For the contact conditions studied (contact pressure $1.5 \times 10^6 \text{ Pa}$, sliding velocity 0.2 m s^{-1}) there is a factor of about 3.5 difference between the maximum and minimum values of these wear rates. Between the two extremes there is a gradual decrease in wear as the amount of rotation is increased. When the amount of rotation is further increased beyond the minimum point the rate of wear increases. We have also observed a significant difference in the nature of the transferred film and the wear debris in these experiments. When the wear is at its maximum the transferred film is thin and the wear debris is highly fibrillar in appearance. At the minimum rate of wear the transferred film is thicker and the wear debris is in the form of the plates. In contrast the rate of wear and the appearance of the transferred layers and wear debris are insensitive to the introduction of imposed rotation for LDPE in similar conditions. These results confirm that the unusual behaviour of PTFE in the simple pin-on-disk experiment probably arises from the introduction of apparent rotation.

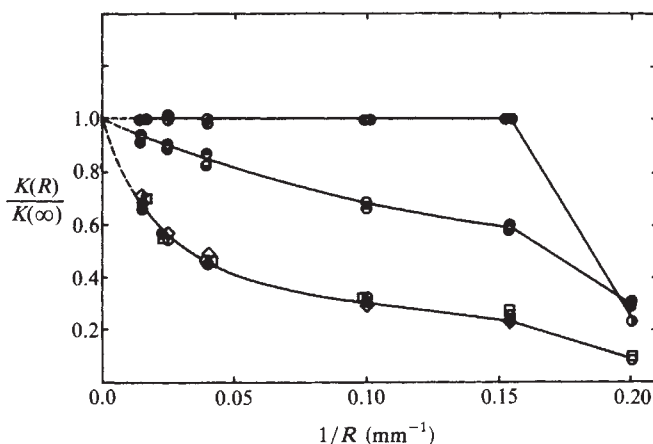


Fig. 2 Normalised wear rate, expressed as $K(R)/K(\infty)$ against $1/R$ where $K(\infty)$ is the estimated wear rate for linear sliding and R the wear track radius. Sliding conditions as for Fig. 1. $\square$, Carbon-filled PTFE 1: 5 w/w; $\bullet$, UHMPE; $\bullet$, PMMA; $\bullet$, HDPE; test pin diameters, 13 mm. Other symbols as for Fig. 1. Polymers of the type HDPE, UHMPE and PTFE show a marked reduction in wear as R decreases, while PMMA and LDPE show no effect until R approaches the radius of the test specimen. For PTFE the reduction in wear is of the order of 70% while for HDPE and UHMPE the effect is more modest, approximately 40%. Further reductions are noted when the wear track radius decreases below that of the radius of the pin.

They further emphasise the peculiar sensitivity of this polymer to rotation during sliding and indicate that these effects arise from the disruption of natural orientation processes which are produced during sliding.

Our results have two consequences: at a practical level they indicate the need for caution when measuring the friction and wear of PTFE and, to some extent, HDPE and UHMPE and their composites using the pin-on-disk configuration, and from an academic point of view it introduces a new and interesting variable for probing the friction and wear mechanism of these materials.

T.A.S. thanks the British Council for a research fellowship.

Received 2 May; accepted 24 July 1979.

1. Tribol. Int. 6(6), 213–293 (1973).
2. Briscoe, B. J. & Tabor, D. in *Polymer Surfaces* (eds Clark, D. T. & Feast, W. J.) (Wiley, New York, 1978).
3. Lee, L. H. (ed.) *Advances in Polymer Friction and Wear* (Polymer Science and Technology Series 5A and 5B, Plenum, New York, 1974).
4. Brown, R. D. *Test Methods in Boundary Lubrication, An Appraisal of World Literature* (eds Ling, E. F., Klaus, E. F. & Fein, P. S.) (ASME, London, 1969).
5. Lancaster, J. K. in *Polymer Science, A Materials Science Handbook* (ed. Jenkins) Ch. 14 (North-Holland, Amsterdam, 1972).
6. Briscoe, B. J., Steward, M. D. & Groczek, K. A. *Wear* 42, 99–102 (1977).
7. Pooley, C. M. & Tabor, D. *Proc. R. Soc. A* 329, 251–274 (1972).
8. Briscoe, B. J. & Tabor, D. *Self-Lubricating Polythene* (Colloques Internationaux du CNRS, No. 233, 1974).
9. Briscoe, B. J., Pogossian, A. K. & Tabor, D. *Wear* 27, 19–34 (1974).

First observation of the coil-globule transition in a single polymer chain

Izumi Nishio, Shao-Tang Sun, Gerald Swislow & Toyochi Tanaka*

Department of Physics and Center for Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

The conformations of single polymer chains in solution have been studied extensively since the 1940s. At high temperatures and in good solvents, a polymer has an extended coil configuration, while at low temperatures and in poor solvents a polymer is in a collapsed globule state. The transition between the extended and collapsed state as the temperature or solvent composition is varied was thought to be smooth and continuous, and experiments supported this idea¹. However, in the 1960s it was suggested^{2–4} that the transition between the two configurations, the coil-globule transition, was discrete. We now report the first observations of the single polymer coil-globule transition. The observations were made on polyacrylamide molecules dissolved in acetone-water mixtures.

The mechanism of the transition, as described by Lifshitz⁴, is quite similar to that of the gas-liquid transition. The extended polymer corresponds to the gas state, whereas the collapsed polymer corresponds to the liquid state. Many others have contributed to the theory of the coil-globule transition^{5–9}. For example, de Gennes⁹, using the analogy of the magnetic transition, pointed out that the coil-globule transition of an infinitely long polymer may correspond to a tricritical point. On the experimental side, there have been reports on the temperature dependence of the radius of gyration of polystyrene in cyclohexane^{10–13}, and on polyacrylic acid in dioxane¹³. In these experiments, however, the existence of the coil-globule transition was not confirmed.

Recently, Tanaka¹⁴ and Fillmore and Tanaka¹⁵ observed the reversible collapse of a covalently crosslinked polyacrylamide network immersed in an acetone-water mixture. The collapse occurred upon changing either the temperature or acetone

concentration of the system. The experiments showed that the volume of the gel can change reversibly by a factor of 100 at the transition. This suggests that a single polyacrylamide chain may undergo a coil-globule transition in acetone-water mixtures.

Acetone is a poor solvent for polyacrylamide, whereas water is a good solvent for the molecule. By gradually changing the composition of the mixture, the polyacrylamide molecules can be brought through the transition between the regions of good solvent and of poor solvent. In our experiment, we used monodisperse polyacrylamide polymers of molecular weight $5-6 \times 10^5$ (Polysciences Inc). A single chain of the polymer consists of approximately 8×10^4 acrylamide monomers and has a length of about $24 \mu\text{m}$. The concentration of the polymer in solution was less than $10 \mu\text{g ml}^{-1}$. At this concentration, the mean distance between adjacent polymers, of the order of $1 \mu\text{m}$, was much larger than the average polymer size, thus inter-polymer entanglements were avoided. An argon-ion laser (Spectra Physics Model 165) at $5,145 \text{ \AA}$ wavelength was used as a light source. The laser power was approximately 1 W. The correlation function of light scattered from the sample at a 90° angle was measured using a photomultiplier tube and a photon correlator (Nicoli Instruments). All measurements were made at 25°C . From the decay rate of the correlation function of the scattered light intensity, we determined the diffusion coefficient, D , of the polymer, which is related to the hydrodynamic radius, ξ , of the polymer through the Stokes-Einstein relation¹⁶:

$$D = kT/(6\pi\eta\xi)$$

where k is the Boltzmann constant, T is the absolute temperature, and η is the viscosity of the solvent.

Figure 1 shows the hydrodynamic radius of the polymer plotted as a function of acetone concentration (open circles). At low acetone concentrations, the radius is large, approximately $500 \text{ \AA}$. Near an acetone concentration of 39%, the polymer shows a drastic decrease in hydrodynamic radius to about

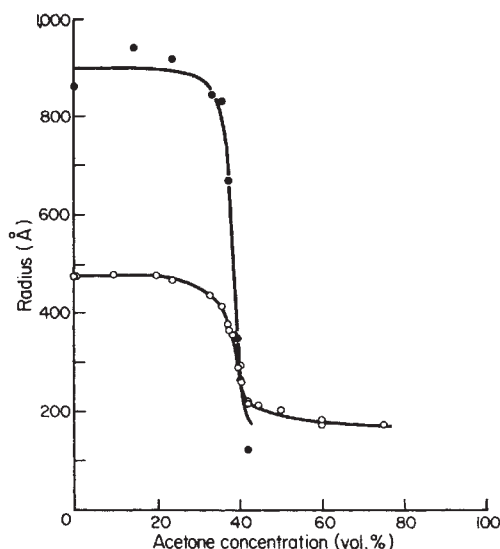


Fig. 1 The hydrodynamic radius of a single polyacrylamide chain, determined by laser light scattering spectroscopy, is plotted (O) as a function of the acetone concentration of the acetone-water mixtures used as solvent. The radius of gyration of the polyacrylamide chain, determined from dissymmetry measurements of the scattered light intensity, is also plotted (●). It was very difficult to determine the radius of gyration for values below $300 \text{ \AA}$. Samples were prepared as follows. Polyacrylamide molecules were carefully dissolved in pure water, and gently stirred using a magnetic stirrer. The solution was then left standing for several days to ensure complete dissolution, before adding the solution to the acetone-water mixtures. The samples were then filtered through a $0.4\text{-}\mu\text{m}$ filter before the measurements were made.

* To whom reprint requests should be addressed.

200 Å. With further increases in acetone concentration, the polymer radius remains constant. We found the transition to be reversible and reproducible.

We also determined the radius of gyration of the chain from the angular dissymmetry of the scattered light intensity¹⁷. The results are shown in Fig. 1 (solid circles). The coil-globule transition is seen to occur at exactly the same acetone concentration as that seen in the curve of the hydrodynamic radius. The value for the radius of gyration, r_g , is about twice that of the hydrodynamic radius, ξ , in the coil state for the following reasons. First, for chain polymers, the hydrodynamic radius may be smaller than the radius of gyration due to such factors as the nature of the side groups on the polymer chain and the viscosity of the solvent. Measurements on polystyrene in solution¹⁸ have shown ξ to be at least 20% smaller than r_g . The difference would be expected to be even larger for polyacrylamide, as it has much smaller side groups. Second, in our experiments, when r_g becomes higher than 500 Å, the fluctuations in the chain conformation begin to scatter light substantially^{19,20}. This makes the apparent diffusion coefficient, calculated assuming that the scattering light intensity fluctuates due only to the translational motion of the polymer, larger than its actual value. Thus, the apparent ξ becomes less than the actual value. It is of interest to investigate further the amplitude and relaxation times of such chain configuration fluctuations.

The observed transition between the coil and globule states of the polymer is quite sharp, but not absolutely discrete. The finite transition width is probably due to the polymer chain comprising only a finite number of monomer units. A chain composed of an infinite number of monomers would be required to observe a

discrete transition⁴. It will be important to investigate further the effect the finite size of the system has on the transition.

Thus we have presented direct evidence of the existence of the coil-globule transition in single polymers. The hydrodynamic radius of polyacrylamide chains in solution changes sharply in the transition between the regions of good solvent and of poor solvent. This sharp change in polymer size may function in amplification, switching, and memory at the molecular level and could have potential applications in biology, medicine and chemical engineering.

Received 4 June; accepted 6 August 1979.

1. Flory, P. J. *Principles of Polymer Chemistry*. (Cornell University Press, Ithaca, 1953).
2. Ptitsyn, O. B. & Eizner, Y. Y. *Biofizika* **10**, 1-5 (1965).
3. Eizner, Y. Y. *Polymer Sci. USSR* **11**, 409-417 (1969); **14**, 1695-1706 (1972).
4. Lifshitz, I. M. *Ah. Eksp. Theor. Fiz.* **55**, 2408-2422 (1968).
5. Edwards, S. F. in *Critical Phenomena* (eds Green, M. S. & Sengers, J. V.).
6. Domb, C. *Polymer* **15**, 259-262 (1974).
7. Lifshitz, I. M., Grosberg, A. Y. & Khokhlov, Rev. mod. Phys. **50**, 683-714 (1978).
8. De Gennes, P. G. *J. phys. Lett.* **39**, 299-301 (1978).
9. De Gennes, P. G. *J. phys. Lett.* **L36**, 55-57 (1975).
10. Mazur, J. & McIntyre, D. *Macromolecules* **8**, 464-476 (1975).
11. Mashimo et al. *J. Phys. Soc. Jap.* **39**, 1529-1537 (1975).
12. Nierlich, M., Cotton, J. P. & Farnous, B. *J. chem. Phys.* **69**, 1379-1383 (1978).
13. Cuniberti, C. & Bianchi, U. *Polymers* **15**, 346-350 (1974).
14. Tanaka, T. *Phys. Rev. Lett.* **40**, 820-823 (1978).
15. Fillmore, D. J. & Tanaka, T. (in preparation).
16. Benedek, G. B. in *Polarization, Matter and Radiation*, 49-84 (Presses Universitaires de France, Paris, 1969).
17. Doty, P. & Edsall, J. T. *Advances in Protein Chemistry* Vol. 6, 35-121 (Academic, New York, 1951).
18. Nishio, I. & Wada, A. (in preparation).
19. Pecora, R. *J. chem. Phys.* **43**, 1562 (1965).
20. Adam, M. & Delsanti, M. *Macromolecules* **10**, 1229-1237 (1977).

Basaltic pillars in collapsed lava-pools on the deep ocean-floor

J. Francheteau*, T. Juteau† & C. Rangan‡

* Centre Océanologique de Bretagne, BP 337, 29273, Brest Cedex

† Université Louis Pasteur, Strasbourg

‡ Université Pierre et Marie Curie, Paris, France

Observations of peculiar volcanic objects, made by a submersible on the deep sea floor at a depth of about 2,600 m at and near the axis of the East Pacific Rise during the CYAMEX expedition as part of the RITA Project¹ are presented here. Two basic types of flow forms were observed within the crestal area of the East Pacific Rise: pillow flows and fluid lavas, the latter sometimes overlying massive flows. The East Pacific Rise at 21° N comprises an axial, unfaulted extrusion zone bordered by an extension zone characterised by faulting^{2,3}. Pillow flows occupy the innermost or extrusion zone and constitute small elongated volcanic highs. Fluid lavas tend to occur at the edge of the adjacent extension zone in bathymetric lows controlled by normal faults or steep primary slopes of constructional highs. In the 50 × 200 m lows which border the extrusion zone the fluid lava is smooth and lobate surfaces which represent the upper surface of the flow are locally collapsed and reveal the internal structure of the fluid lavas. Where the roof collapse is extensive, layered columnar features are visible and volcanic layering can be seen against the flank of the bordering volcanic highs (Figs 1-3). Similar features have been reported from the Galapagos Rift⁴. The diameter of the approximately cylindrical pillars ranges from 0.5 to 2 m. Some pillars are made of multiple coalescent cylinders. The tops of the pillars are glassy, funnel-shaped and always widening upwards. The pillars were presumed to be hollow from several observations of gashes or openings in the vertical walls of the pillars. This was demonstrated during dive CY 78-19 to the south where a small pillar

was toppled by CYANA and subsequent examination revealed a circular canal along the axis of the pillar. The outer surface of the pillars is marked by centimetre-thick glassy, subhorizontal ledges extending several centimetres from the outer vertical surface of the pillars (Figs 2-4). The ledges are spaced every 2-5 cm and show small lava stalactites hanging on the underside of the ledges. Examination of large layered fragments of pillars recovered by CYANA demonstrated that the layering is only a surface feature as it does not extend through the basaltic mass of the pillars. The apparent layering is due to glass ledges adhering to a vertical basaltic pipe. In some rare instances the pillar outer surface showed no ledges and instead a smooth surface corrugated with vertical grooves. Some pillars are inclined or slightly curved; others get narrower towards the base. The pillars are almost totally aphyric and have the same bulk composition as other lava types recovered in the axial zone of the East Pacific Rise at 21° N (ref. 5).



Fig. 1 General landscape: pillars within collapsed lava pond.



Fig. 2 Detail of a pillar showing the centimetric pseudo layering, with thin and darker salient glass layers projecting from the basaltic surface of the pillar.

The pillars occur in forest-like arrangement (Fig. 3). Their spacing ranges from 0.5 to 3 m and sometimes they coal c, building walls, especially near the edges of the depressions. However, the pillars are not confined to the edges. The flanks of the lows show large segments of undisrupted glass ledges analogous to those observed on the pillar surfaces. In one case (dive CY 78-19) where the fluid flows partially filled a graben, the layered lateral margin could be seen to represent a chilled, 5-10 cm-thick wall plated against the truncated pillow basalts of the normal fault scarp. The chilled margin was partially detached from the scarp surface.

Non-collapsed glassy upper surfaces of fluid lavas surrounding the fields of pillars show that in the depressions, the thickness of the lobate fluid lava roof is about 10 cm. The horizontal dimension of the lobes is of the same order as the pillar spacing and pillars provide roof support at the junction between adjacent lobes. The geometry of the lobate roof junction is analogous to that of the funnel-shaped pillar tops. Pieces of the lobate roof are commonly preserved around the fluid lava depressions and also where they form bridges between adjacent pillars. In most cases the lava pool appears to be concave upward with a slight (a few metres) depression in the middle of the pool with respect to the edges.

The seafloor around the pillars is strewn by rubble made up of pillar fragments and roof slabs. A deeply incised rille-like canyon was observed near the axis of the fluid-lava depression.

These observations lead us to the following interpretations (Figs 5 and 6):

(1) The fluid lavas fill a preexisting depression either controlled by faulting or by adjacent constructional highs. This is proved by inspection of the lateral contact and by the contrast in lava freshness between the fill and the adjacent pillows. The remarkable fluidity of the lavas cannot be ascribed to chemical differences or to temperature of extrusion. It probably reflects the large volume of outpouring which could in turn be a function of a very high spreading rate over a short time span.

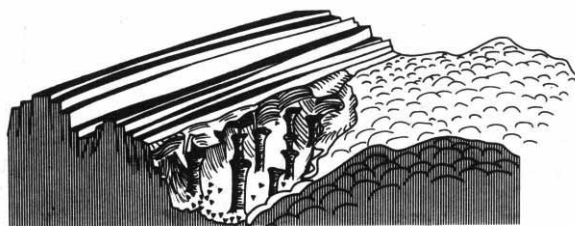


Fig. 3 Lava pool after roof collapse showing basaltic pillars at the boundary of the unfractured extrusion zone (on the right). The pillars are about 10-15 m high.

(2) The seawater which is trapped in the crust under the fluid lava pool is heated and expands (by a factor of 20 for a temperature of 800 °C at a depth of 2,600 m (see ref. 6), forcing its way up through the molten lava in a series of vertical conduits (J. Moore, personal communication) (Fig. 5). The walls of the conduits will be quenched at the boundary between the rising water and the hotter magma. The conduit walls will grow outward from the inner glassy walls of the conduits with time. This explains the hollow pillars found by CYANA. A similar explanation has been put forward by Fuller⁷ and Waters⁸ to account for the existence of vertical cavities (spiracles) that cut massive lava on the Columbia River Plateau. These authors proposed that water vapour and other gases surge upward into liquid lava that covers marshy ground.

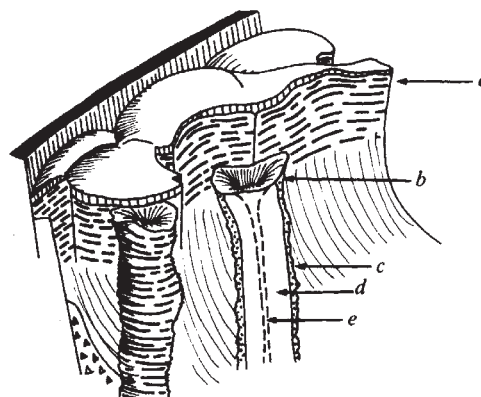


Fig. 4 Detail of basaltic pillar and of edge of magma pool. *a*, Pillar attached to the edge of the pool and capped by smooth lobate fluid lava. *b*, Pillar: funnel-shaped widening upwards. Pillar cross-section showing: *c*, lava cooling ledges on outer margin; *d*, massive lava inside pillar wall; *e*, hollow conduit inside pillar.

(3) Before any extensive crystallisation takes place, the lava pool is drained rapidly in successive stages (Fig. 6). This leaves 'bath-tub' rings around the 'cold walls' present in the depression. This includes the pillars and the edges of the pool. This accounts for the apparent layering shown by the pillar samples. The rille-like canyon seen on the floor of the pool may be a drain path. Finally down-drag of either the roof or the outer rigid carapace of the pillars can scrape vertical grooves on the still plastic interior and explain the smooth corrugated outer surfaces of some pillars. A similar mechanism has been proposed by Moore and Richter⁹ to explain the vertical striations on the surface of Hawaiian tree moulds.

(4) Solidification on the fluid lava against the cold wall at the bottom of the pool would help to maintain pillars in position after collapse of the roof.

Thus tall, near vertical and hollow pillars with rings of glassy ledges are found in dense arrangement within fossil pools of fluid lava. They are interpreted as the fossil witnesses of both

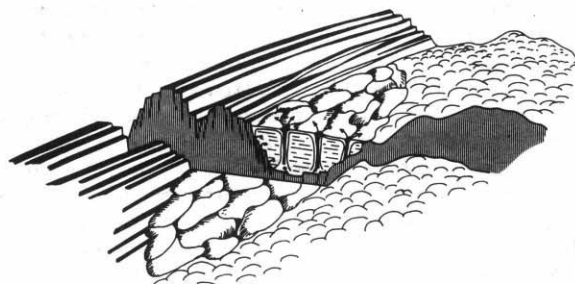


Fig. 5 Reconstruction of magma pool before the draining of lavas and roof collapse. The arrows at the pool surface represent escape of water.

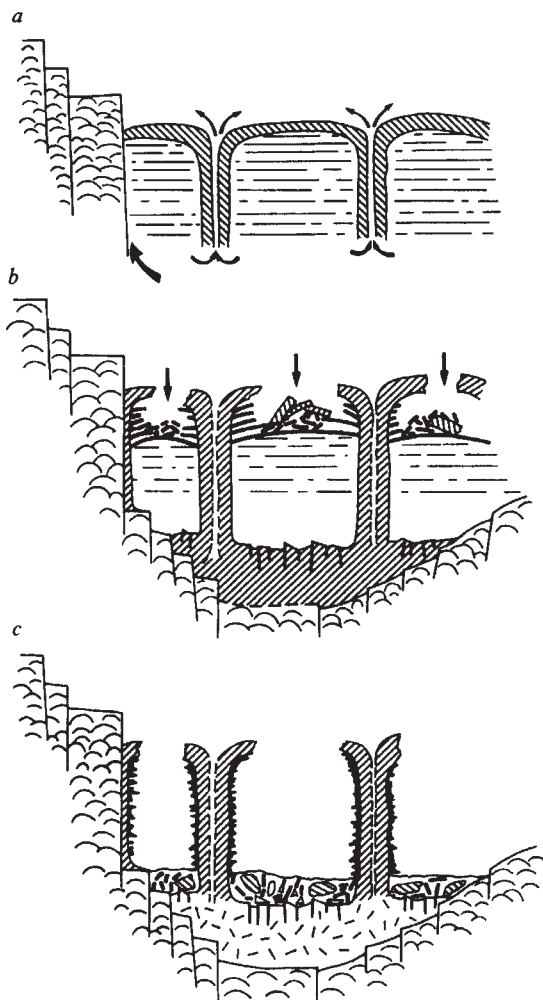


Fig. 6 Inferred evolution of magma pool showing history of pillars and cooling ledges at several stages (see text). *a*, Initial stage of filling of the depression; *b*, beginning of lava pool withdrawal with concomitant roof collapse. *c*, final stage of roof collapse showing isolated pillars standing among coarse rubble.

temporary liquid magma pools which can fill topographic lows and smooth the ridge topography, and of the rhythmic nature and rapidity of the magma pool withdrawal. When drain-off is prevented or when it proceeds too slowly, massive lavas are created.

The fluidity of the lava can be explained if there is rapid emplacement of a large volume of lava resulting in slow cooling of the bulk of the lava pool. The large volume of outpouring may be due to a high 'instantaneous' spreading rate concomitant with a high ascent rate of the lava.

The observations at 21° N show that the size of lava pools increases southward along the axis and thus that the fluid lava regime in the south is predominant over the pillow lava regime. This suggests that instantaneous rates of spreading may be variable along the strike of the mid-ocean ridge.

Received 27 April; accepted 30 July 1979.

1. CYAMAX Scientific Team *Eos* **59**, 1198 (1978).
2. Normark, W. R. *Geology* **4**, 681–685 (1976).
3. CYAMEX Team, Bougault, H., Cambon, P. & Eissen, J. P. Submarine basalts collected by diving saucer on the East Pacific Rise at 21° N, and on the Tamayo Transform Fault (in the press).
4. Ballard, R. D., Holcomb, R. T. & Van Andel, T. H. Sheet Flows, Collapse Pits and Lava Lakes of the Rift Valley (in the press).
5. CYAMEX Scientific Team Manned submersible dives on the East Pacific Rise 21° N (Project RITA) (in the press).
6. McBirney, A. R. *Bull. Volcanol.* **26**, 455–469 (1963).
7. Fuller, R. E. *Am. J. Sci.* **21**, 281–300 (1931).
8. Waters, A. C. *Am. J. Sci.* **258-A**, 350–366 (1960).
9. Moore, J. G. & Richter, D. H. *Bull. geol. Soc. Am.* **73**, 1153–1158 (1962).

Planktonic foraminiferal fauna associated with eastern Mediterranean Quaternary stagnations

Robert C. Thunell

Department of Geology, University of South Carolina, Columbia, South Carolina 29208

G. P. Lohmann

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543

Bradley¹ predicted, before any evidence from sediment coring in the Mediterranean Sea, that changing late Quaternary climatic conditions would result in periodic stagnations of the eastern basin and with them, accumulations of organic-rich sediments. Such organic-rich layers from the eastern Mediterranean were subsequently observed in sediment cores by Kullenberg² and later named "sapropelitic layers"³. Since these initial observations, sapropels have been reported in most studies of Quaternary sediments from the Ionian Sea^{4–14}. As a result, there is a great deal of information available concerning the chemical^{15–18} and sedimentological properties^{13–14}, faunal characteristics¹⁹ and temporal^{9–11} and spatial^{10,12} distribution of late Quaternary sapropels. In the study reported here we have attempted to characterise the planktonic foraminiferal biogeography of such late Quaternary sapropels in terms of the modern distribution of foraminifera in the Mediterranean Sea.

Sapropels are black, organic-rich muds with unusually high concentrations of hydrogen sulphide and iron hydroxide¹⁵. These sediments seem to be deposited in a particular sequence^{13–14}, reflecting the transition from open-marine sedimentation (hemipelagic) to that indicative of stagnant conditions. Sapropels are found throughout the entire eastern Mediterranean at depths greater than 800–1,000 m (ref. 12), although the areal extent of individual layers may not be basin-wide¹⁰.

Distinct microfaunal and microfloral assemblages have been described from late Quaternary sapropels. Planktonic foraminiferal^{5,9,10,19,21}, nannoplankton²⁰, diatom¹⁶ and pteropod⁴ assemblages in sapropels are indicative of reduced surface salinities. For example, high frequencies of *Neogloboquadrina dutertrei*, a planktonic foraminifer preferring low surface salinities^{22,24}, are generally associated with late Quaternary sapropels^{2,5,9–11,21}. Benthonic Foraminifera and ostracodes, as well as any evidence of bioturbation, are totally absent from these layers^{4,5,10}.

Quaternary glacial-interglacial fluctuations have been a fundamental element of even the earliest theories of sapropel formation. Both Bradley¹ and Kullenberg² suggested that stagnation of the eastern Mediterranean resulted from lowered surface salinities produced by increased precipitation and runoff during glacial times. Subsequent studies^{3,5,7,8,10,11} have also invoked a low salinity surface layer during sapropel formation, but demonstrated that this occurred during interglacial episodes. According to these studies, the flushing of freshwater from the Black Sea into the eastern Mediterranean, possibly combined with increased precipitation, forms a low salinity surface layer that prevents oxygenation of the deep basins and results in accumulation of anoxic sediments. The previously mentioned faunal evidence, together with recent studies on the oxygen isotopic composition of planktonic Foraminifera from sapropel layers, strongly support the existence of a low salinity surface layer^{11,19,25}.

The present study involves three basic steps: (1) characterisation of the sapropel-associated faunas; (2) comparison of

these sapropel-associated faunas with the modern distribution of planktonic foraminifera; and (3) interpretation of the ancient sapropel environment in terms of the modern oceanographic conditions where faunas similar to those associated with sapropels are found. We have simply tried to find out where faunas most similar to the sapropel assemblage live in the modern Mediterranean and determine what is oceanographically unique about these environments.

The sapropel-associated planktonic foraminiferal assemblage was determined by averaging the faunas found in 12 late Quaternary sapropels from cores TR171-27 (33°50' N, 25°59' E) and TR172-22 (35°19' N, 29°01' E). The averaging process involves a *Q*-mode principal components analysis of all the foraminiferal frequency data obtained for the sapropel layers. The resulting first principal component (first eigenvector of the pseudo $\cos \theta$ matrix²⁶) represents the 'average sapropel fauna' for the late Quaternary in this region (Table 1). It accounts for ~74% of the variance in the original data, with four species being particularly important (*Globigerinoides ruber*, *Neogloboquadrina dutertrei*, *Globigerina bulloides* and *Orbulina universa*). These species have been previously shown to generally increase in abundance within sapropel layers^{5,7,9,10,21,27}.

The present-day area distribution of this average sapropel fauna was found by comparing its composition with the planktonic foraminiferal compositions of 66 core top samples distributed throughout the Mediterranean. The core top faunal data, as well as other pertinent information on these samples, has been published previously²⁸. The measure of similarity used is $\cos \theta$, where θ is the angle between the first principal component and a given sample. It is calculated simply as the product of the unit sample vectors and the unit eigenvector, and it varies from 0 to 1.

The area where core-top faunas are most similar to the average late Quaternary sapropel assemblage is shown in Fig. 1a (where similarities are greater than 0.8 on a $\cos \theta$ scale of 0–1.0).

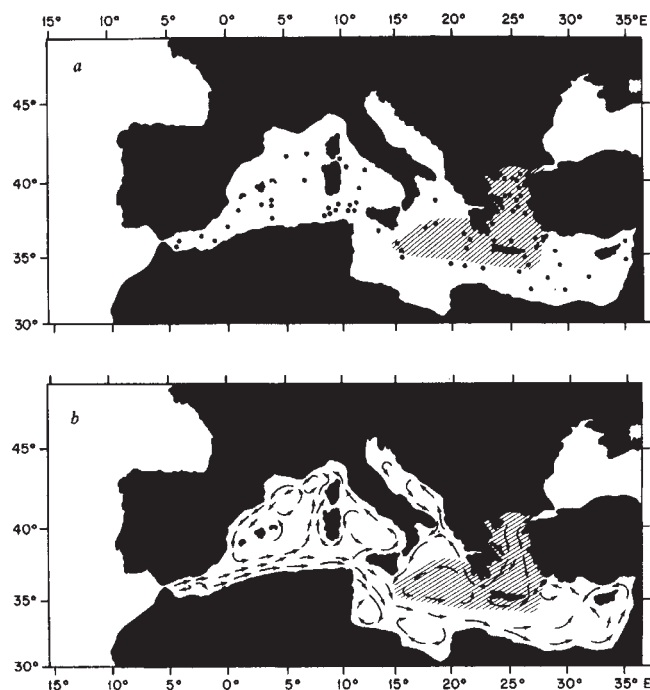


Fig. 1 *a*, Area distribution of the modern planktonic foraminiferal fauna most similar to the average sapropel fauna (region where similarities are greater than 0.8 on a $\cos \theta$ scale of 0–1.0 is indicated). The core top faunal data is from ref. 28. *b*, Present day surface circulation of the Mediterranean presented by Lacombe and Tchernia³⁰. The region containing the modern planktonic foraminiferal fauna most similar to the average sapropel fauna is indicated.

Table 1 Average sapropel fauna derived from a *Q*-mode principal components analysis of 18 planktonic foraminiferal species from 12 late Quaternary sapropel layers

Species	Per cent abundance
<i>Globorotalia inflata</i>	0.7
<i>Globorotalia truncatulinoides</i> (dextral)	0.0
<i>Globorotalia truncatulinoides</i> (sinistral)	0.5
<i>Globorotalia scitula</i>	1.6
<i>Globigerinoides ruber</i>	29.8
<i>Globigerinoides sacculifera</i>	1.1
<i>Globigerinoides tenellus</i>	1.3
<i>Globigerina bulloides</i>	18.6
<i>Globigerina calida</i>	0.0
<i>Globigerina digitata</i>	0.8
<i>Globigerina falconensis</i>	0.9
<i>Globigerina quinqueloba</i>	1.9
<i>Globigerinella aequilateralis</i>	3.6
<i>Globigerinita glutinata</i>	4.0
<i>Neogloboquadrina dutertrei</i>	24.3
<i>Neogloboquadrina pachyderma</i> (dextral)	1.9
<i>Neogloboquadrina pachyderma</i> (sinistral)	0.2
<i>Orbulina universa</i>	8.8

Abundances are loadings of species on the first eigenvector of the pseudo $\cos \theta$ matrix²⁶ and are converted to per cent.

This region of maximum similarity includes the Aegean Sea and the area immediately west of Crete. A comparison of the area distribution of the modern fauna most similar to the average sapropel fauna with the present-day surface circulation of the Mediterranean reveals a close association with surface water that flows out of the Aegean Sea and forms a small gyre in the Ionian Sea (Fig. 1b). At present, the Aegean Sea receives a great deal of low salinity water from both river runoff and the Black Sea, with the salinity of the Black Sea water being ~16‰ when it enters the Bosphorus²⁸. Although this low salinity water mixes rapidly within the Aegean, it is a component of the surface water that flows out of the Aegean. It is within the bounds of this Aegean-derived surface water that the modern planktonic foraminiferal fauna most similar to the average sapropel fauna is confined (Fig. 1a).

This biogeographic pattern may give further insight into the mechanism of sapropel formation. If the late Quaternary stagnations of the eastern Mediterranean are due primarily to the input of fresh water from the Black Sea, one would expect to find a low salinity lens extending from the Aegean over the remainder of the eastern basin during times of sapropel deposition. It seems that the biogeographic pattern resulting from a comparison of the modern fauna with the average sapropel fauna is simply a less severe version of what one would expect during periods of stagnation. That is, the present-day surface flow patterns in this region are analogous to those which were associated with times of stagnation. The fundamental difference between the two regimes is that during sapropel formation an intensification of this pattern occurs due to enhanced freshwater input. These results further support the idea that the Black Sea is the major source of freshwater needed to cause stagnation^{3,5,19}.

We thank Drs W. Berggren, B. Corliss and D. Ross for reviewing the manuscript. Financial support was provided by a Woods Hole Oceanographic Institution postdoctoral fellowship and by NSF grant OCE77-82149 (CENOP).

Received 18 May; accepted 18 July 1979.

- Bradley, W. H. *Science* **88**, 376–379 (1938).
- Kullenberg, B. *Medd. Oceanogr. Inst. Göteborg* **21**, 1–38 (1952).
- Olausson, E. *Rep. Swed. Deep-Sea Exped.* **8**, 353–391 (1961).
- Herman, Y. *The Mediterranean Sea: A Natural Sedimentation Laboratory* (ed. Stanley, D. J.) 129–147 (Dowden, Hutchinson and Ross, Stroudsburg, 1972).
- Ryan, W. B. F. *The Mediterranean Sea: A Natural Sedimentation Laboratory* (ed. Stanley, D. J.) 149–169 (Dowden, Hutchinson and Ross, Stroudsburg, 1972).
- Van Straaten, L. M. *The Mediterranean Sea: A Natural Sedimentation Laboratory* (ed. Stanley, D. J.) 631–643 (Dowden, Hutchinson and Ross, Stroudsburg, 1972).
- Cita, M. B. *et al. Init. Rep. DSDP Leg 13*, 1263–1339 (1973).
- McCoy, F. W. thesis, Harvard Univ. (1974).

9. Cita M. B. *et al. Quat. Res.* **8**, 105–235 (1974).
10. Thunell, R. C., Williams, D. F. & Kennett, J. P. *Mar. Micropaleont.* **2**, 371–388 (1977).
11. Vergnaud-Grazzini, C., Ryan, W. B. F. & Cita, M. B. *Mar. Micropaleont.* **2**, 371–388 (1977).
12. Stanley, D. J. *Nature* **274**, 149–152 (1978).
13. Maldonado, A. & Stanley, D. J. *Mar. Geol.* **20**, 27–40 (1976).
14. Maldonado, A. & Stanley, D. J. *Geology* **5**, 111–117 (1977).
15. Nesteroff, W. D. *Init. Rep. DSDP Leg 13*, 713–720 (1973).
16. Rossignol-Strick, M. *Init. Rep. DSDP Leg 13*, 971–991 (1973).
17. Deroo, G., Herbin, J. P. & Roucace, J. *Init. Rep. DSDP Leg 42A*, 465–472 (1978).
18. Sigl, W., Chamely, H., Fabricius, F., d'Argoud, G. & Muller, J. *Init. Rep. DSDP Leg 42A*, 445–465 (1978).
19. Williams, D. F. & Thunell, R. C. *10th Int. Congr. Sedimentology*, Jerusalem, 738–739 (1978).
20. Muller, C. *Bull. geol. Soc. Greece* **10**, 133–144 (1973).
21. Kidd, R. B., Cita, M. B. & Ryan, W. B. F. *Init. Rep. DSDP Leg 42A*, 421–443 (1978).
22. Ruddiman, W. F. *Bull. geol. Soc. Am.* **82**, 283–302 (1971).
23. Be, A. W. H. & Tolderlund, D. S. *The Micropaleontology of the Oceans* (eds Funnell, B. M. & Reidel, W. R.) 105–149 (Cambridge University Press, London, 1971).
24. Kipp, N. G. *Mem. geol. Soc. Am.* **145**, 3–41 (1976).
25. Williams, D. F., Thunell, R. C. & Kennett, J. P. *Science* **201**, 252–254 (1978).
26. Klován, J. E. & Imbrie, J. J. *Int. Ass. Math. Geol.* **3**, 61–77 (1971).
27. Williams, D. F. & Thunell, R. C. *Sed. Geology* (in the press).
28. Thunell, R. C. *Mar. Micropaleont.* **3**, 147–173 (1978).
29. Sverdrup, H., Johnson, M. & Flemming, R. *The Oceans* (Prentice Hall, New Jersey, 1942).
30. Lacombe, H. & Tchernia, P. *The Mediterranean Sea: A Natural Sedimentation Laboratory* (ed. Stanley, D. J.) 25–36 (Dowden, Hutchinson and Ross, Stroudsburg, 1972).



Fig. 1 Occlusal view of PUA15 and PUA16 ($\times 7$). (The characteristic features of the molar crowns are as identified in Fig. 2a.)

Miocene tree shrews from the Indian Sivaliks

S. R. K. Chopra, S. Kaul & R. N. Vasishat

Department of Anthropology, Panjab University,
Chandigarh-160014, India

There have been no definite reports of fossil tree shrews although a rib cage from the Tatrot Formation of India has been identified as *Tupaia*¹, and fossil material, including the anterior part of a skull (YGSP 8089) with broken crowns of teeth (I^1 – P^4) and a lower left molar (YGSP 8090), from the Middle Sivaliks of Pakistan, has been assigned to the sub-family Tupaiinae (unpublished results by I. Jacobs). The lack of material makes it difficult to assess the evolutionary position of this group, which, together with moles and hedgehogs, is considered to include the most direct modern descendents of the primitive placentals². The living tree shrews have also attracted great interest because of the suggestion that they are related to primates. We report here a maxillary fragment (PUA 15) of the left side with M^1 – M^3 and an isolated right lower molar, probably M^2 , (PUA 16) recovered from the middle Sivalik deposits of Haritalyangar in Himachal Pradesh, India. We have referred this to the Tupaiinae, although we cannot assign generic and specific names. The following communication³ describes more substantial material discovered later.

The maxillary fragment (Fig. 1a) consists of the molars M^1 – M^3 and the posterior roots of premolar P^4 . The molars, with the exception of the third, are quadrate in shape and decrease in size from front to back. Each tooth has three prominent cusps—paracone, metacone and protocone—which are connected by crests. The paracone and metacone are V-shaped, well-separated and near the middle of the tooth. Connected by the centrocrista comprising postparacrista and premetacrista, these two cusps form a W-shaped outer (buccal) wall. The distolingual aspect of M^1 has a small hypocone which is larger than that on M^2 . The hypocone is low and near the neck of the tooth. The buccal portion of the crown of M^3 is broken, but in general this tooth is similar to the first two molars though much smaller. The metacone is also smaller and there is no hypocone on this tooth.

Buccal aspects of M^1 and M^2 display three small styles—parastyle, mesostyle and metastyle. The mesostyle on M^2 is larger than that on M^1 . Because the roots are not visible in the maxillary fragment, nothing definite can be said about their nature and disposition. The wear on the occlusal surfaces is considerable.

The isolated lower second molar of the right side (Fig. 1b) has the basic tribosphenic pattern with an elevated trigonid and a low talonid. Three prominent cusps—paraconid, metaconid and protoconid—constitute the trigonid. The talonid basin, which is larger than the trigonid, is bordered by the hypoconid, hypoconulid and entoconid. A moderately developed precingulum is present on the mesial aspect of the tooth. It is supported by two strong roots; the anterior one is broken almost halfway. The tooth shows slight wear due to use.

The maxillary fragment and the isolated M^2 (Fig. 1) have notable general similarities to those of the Tupaiinae (Fig. 2a). Like them, the upper molars lack internal cingula on the mesial

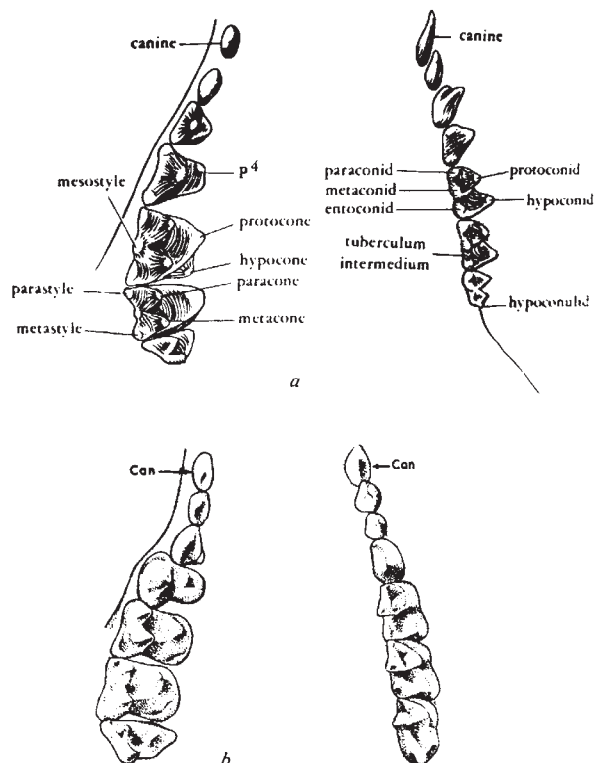


Fig. 2 Occlusal view of the left upper (left) and lower (right) teeth of *Tupaia glis* (a, $\times 2.2$) and *Ptilocercus* (b, $\times 3.4$). (a After Swindler⁸, b after Clarke⁴.)

and distal aspects. The buccal cingulum is visible but the lingual cingulum is not. There is no tuberculum intermedium arising from the entoconid as it does in the lower molars of *Tupaia glis* (Fig. 2a). Like most members of the Tupaiidae, and unlike the Soricidae⁵, M¹ and M² have distinct hypocones. Unlike those of *Ptilocercus* (Fig. 2b), the molars have a mesostyle and are not surrounded by a distinct cingulum on all sides⁶⁻⁸.

A plausible taxonomic placing would need a more complete dentition than is available. However, similarities in morphology with the members of the Tupaiidae outweigh the differences, and, without assigning either generic or specific names at present, we believe that this material warrants its position in the sub-family Tupaiinae.

We thank Baldev Singh for help with photographs.

1. Dutta, A. K. *India Minen*, 29, 76 (1975).
2. Romer, A. S. *Vertebrate Paleontology* (University of Chicago Press, 1966).
3. Chopra, S. R. K. & Vasishat, R. N. *Nature* 281, 214-215 (1979).
4. Clark, W. E. Le Gros. *The Antecedents of Man* (Edinburgh University Press, 1959).
5. Gromova, V. I. (ed.). *Fundamentals of Paleontology*, 13 (Israel Program of Scientific Translations, Jerusalem, 1968).
6. Lyon, M. W. *Proc. U.S. natn. Mus.* 45, 1 (1913).
7. Steele, D. G. *Symp. Fourth Int. Cong. Primat.* 3, 154 (1973).
8. Swindler, D. R. *Dentition of Living Primates* (Academic, London, 1976).

Śivalik fossil tree shrew from Haritalyangar, India

S. R. K. Chopra & R. N. Vasishat

Department of Anthropology, Punjab University, Chandigarh-160014, India

We report here a new taxon of fossil tree-shrew from the Śivalik deposits of Haritalyangar in Himachal Pradesh, India. There are no definitive records of fossil tree-shrews, for although two forms, *Anagale*¹ and *Adapisoriculus*² have been reported, their affinities with living tree shrews are disputed. Even the taxonomic position of the tree shrews is controversial, with some workers classifying them as primates³⁻¹⁰ and others disagreeing. Butler¹¹ has suggested that they should be classified in a new order, Scandentia. The skull fragment (PUA I-3) described here is better preserved and much more complete than previous finds, and it has helped us to extend and confirm the identification of the specimens reported in the preceding communication¹².

Significant contributions to the understanding of anatomy, particularly of the skull and dentition of modern tree shrews have come from Clark Le Gros^{5,6}, Davis¹³, Lyon¹⁴ and Steele¹⁵, and we have adopted the classification of Simpson¹⁶ as follows:

Order	Primates
Superfamily	Tupaioidea
Family	Tupauidae
Subfamily	Tupaiinae
Genus	<i>Palaeotupaia</i> new genus

Type: *Palaeotupaia sivalicus*, new species. [Named as an ancient form of the extant genus *Tupaia*, and after the Śivalik Hills where the fossil was found.]

Generic diagnosis: This is an ancestral and extinct form of the subfamily Tupaiinae with upper dentition 2.1.3.3. It resembles in some characteristics various members of the Tupaiinae, but resembles more closely *Tupaia* sp. in most anatomical details of skull and dentition. In size it resembles *T. minor*, but differs in having a proportionately longer face, less posterior incisive foramina, distinct protocones on P³, more transverse P⁴ and divided mesostyles. It differs from *Ptilocercus* (subfamily Ptilocercinae) in having a supraorbital foramen, well separated lacrimal and infraorbital foramina, lacrimal foramen in a distinct notch and well within the orbital margin and infraorbital foramen above P³. It also differs in the position of the orbit in relation to the cheek teeth; in having large gaps between the

central incisors—I¹ and I²—and between I² and C, and in having a slight space between C, P² and P³. P³–I¹ is 1.7 × P⁴–M²; there is a large P³; distinct mesostyles are present, and there is no distinct lingual cingulum or hypocone on the molars.

Palaeotupaia sivalicus new species

Type: PUA I-3 is a well preserved partial skull with left and right P³–M² and roots and alveoli of central incisors, left lateral incisor, right canine, right M³ and P². The right lateral incisor is uninterrupted.

Hypodigm: Type plus a maxillary fragment (PUA I-5) of the left side with M¹–M³, and an isolated right lower molar probably M² (PUA I-6) from Haritalyangar, Himachal Pradesh, India. Horizon and locality: One and a half kilometres north-west of Haritalyangar, Bilaspur district, Nagri Formation, Middle Śivaliks.

Specific diagnosis: Same as for the genus.

Description: The holotype (PUA I-3) under description is slightly damaged anteriorly. Posteriorly it is broken behind the level of left M² and behind the roots of M³. Left M² is also chipped along its buccal margins. The palate appears to be fully ossified. The snout is comparatively long. The skull has a large orbit laterally (better preserved on the right side). Most of the cranial sutures are unclear. The maxillo-premaxillary suture arises from the posterior margin of the incisive foramen. The orbital surface of the maxilla is perforated by the apices of the lingual roots of M¹ and M². Ventrally and anteriorly, at the median transverse level of I² and behind the central incisors is a pair of large, oval incisive foramina. Along the lateral margin of the skull and immediately above the mesial root of P³ is the anterior opening of a comparatively medium sized infraorbital foramen. A small foramen is present at the level of P². The lacrimal foramen is in a distinct notch which lies well within the orbital margin. At the dorsal surface of the skull there is a distinct supraorbital foramen, part of which is broken. The central incisors appear to be slightly larger than the lateral incisors and are separated from each other by a large gap. Similarly, I² is separated from I¹ by a comparatively larger gap. The single-rooted canine is set apart from I² by a large diastema. Although the small and single-rooted P⁴ is very near to P³ and C, the gap between the canine and P² is bigger than that between P²



Fig. 1 *Palaeotupaia sivalicus* PUA I-3. a, Ventral view of the skull; b, dorsal view of the skull; c, lateral view of the skull.



Fig. 2 *Tupaia* sp. Ventral view of the skull.

and P^3 . All the teeth from P^3 to M^2 are three-rooted. P^3 and P^4 are bicuspid, the former with a distinct but poorly developed protocone and the latter with quite a prominent and well developed protocone. A complete and fully developed buccal cingulum is also present in P^4 . M^1 and M^2 are triangular and dilambdodontic. Along the buccal margin of M^1 and M^2 are three prominent styles, parastyle, mesostyle, and metastyle. There is no hypocone or lingual cingulum on M^1 and M^2 .

PUA I-3 differs from *Ptilocercus* in having a supraorbital foramen, well separated lacrimal and infraorbital foramina, a single-rooted canine, large P^3 , mesostyles on the molars, no distinct cingulum and hypocones on the molars, and in the size of the incisors position of the orbit in relation to the teeth.

In the subfamily Tupaiinae it differs from *Anathana* which has the orbit (the anterior end) above P^4 rather than M^1 . It can be differentiated from *Urogale* by the position of the orbit (above M^2) and the shorter face. It differs from *Dendrogale* in having a comparatively broader snout, single-rooted canine and a more transverse P^4 . The specimen resembles *Tupaia minor* in size but differs in having a proportionately longer face, less posterior incisive foramen, distinct protocones on P^3 and more transverse P^4 .

On the basis of the differences from various genera of Tupaiinae and the resemblance to *Tupaia* our specimen is clearly closer to *Tupaia*. But in view of morphological differences and the large age gap between our specimens and the living genus it has been assigned to a new genus and species *Palaeotupaia sivalicus*.

The specimen numbers YGSP 8089 and 8090 may also belong to this new taxon. But this can only be confirmed by further detailed studies of the material.

We thank Professor P. M. Butler for expert comments on the identification, and Dr Louis L. Jacobs for useful discussions.

Received 19 April; accepted 3 August 1979.

1. Simpson, G. G. *Am. Mus. Novitates* **505**, 1 (1931).
2. Lemoine, V. *Bull. Soc. géol. France* **13**, 203 (1885).
3. Carlsson, A. *Acta Zool.* **3**, 227 (1922).
4. Le Gros Clark, W. E. *Proc. Zool. Soc., Lond.* **1**, 461 (1924).
5. Le Gros Clark, W. E. *Proc. Zool. Soc., Lond.* **2**, 1053 (1924).
6. Le Gros Clark, W. E. *Proc. Zool. Soc., Lond.* **1**, 559 (1925).
7. Le Gros Clark, W. E. *Proc. Zool. Soc., Lond.* **2**, 1179 (1926).
8. Le Gros Clark, W. E. *Proc. Zool. Soc., Lond.* **2**, 975 (1932).
9. Le Gros Clark, W. E. *The Antecedents of Man* (Edinburgh University Press, 1959).
10. Le Gros Clark, W. E. *The Antecedents of Man*, 3rd edn (Edinburgh University Press, Edinburgh, 1971).
11. Butler, P. M. in *Studies in Vertebrate Evolution* (eds Joyce, K. A. & Kemp, T. S.) 253–265 (Oliver and Boyd, Edinburgh, 1972).
12. Chopra, S. R. K., Kaul, S. & Vasishat, R. N. *Nature* **281**, 213–214 (1979).
13. Davis, D. D. *Field Mus. nat. Hist., Zool. Ser.* **20**, 383 (1938).
14. Lyon Jr, M. W. *Proc. U.S. natn. Mus.* **45**, 1 (1913).
15. Steele, D. G. *Symp. fourth int. Cong. Primatol.* **3**, 154 (1973).
16. Simpson, G. G. *Bull. Am. Mus. nat. Hist.* **85**, 1 (1945).

The significance of beak marks on the wings of an aposematic, distasteful and polymorphic butterfly

D. A. S. Smith

Department of Biology, Eton College, Windsor, Berks, UK

Birds are the principal predators of adult butterflies¹ which they usually hunt on the wing. When successful, they often carry the prey to a perch and attempt to remove the wings with their beaks before eating. If the initial taste is repellent, the butterfly may be allowed to escape more or less uninjured except for a segment (beak mark) removed from its wing. There is evidence that beak-marked specimens are more common (in museum collections) in taxa known to be distasteful or toxic². Based on field data from a population of *Danaus chrysippus* (L.) (Danaiidae), a well known distasteful species³, I have found that large individuals of both sexes are more frequently beak marked than small ones. Comparing two colour morphs (Fig. 1), the rarer one carried more beak marks.

The study was carried out on 235 hectares of grassland with scattered trees at Dar es Salaam, Tanzania. Butterflies were collected with a net, along regular routes, several times a week for 11 consecutive months. Each specimen was marked (to avoid double counting) and the sex, morph, forewing length and beak marks were recorded before release. Relative density was estimated as the number caught per hour.

Of 1,595 butterflies caught, 117 (7.3%) had beak marks. The distribution of beak marks was non-random in two respects. First, beak-marked butterflies were larger ($\bar{x}$ for forewing length = 38.6 mm) than unmarked specimens ($\bar{x}$ = 37.1 mm), $d = 4.9$, $P < 0.001$. Second, of the two morphs studied (Fig. 1), *aegyptius* was more frequently beak-marked (12.7%) than *dorippus* (6.0%), $d = 4.1$, $P < 0.001$. Overall frequency of beak marks (y) was positively correlated with size of prey (x_2) ($r_{xy} = 0.614$, d.f. 9, $P < 0.05$) and frequency of *aegyptius* (x_3) ($r_{xy} = 0.76$, d.f. 9, $P < 0.01$) but not with density (x_1) (Table 1).

Several observers have claimed that birds tend to choose the largest prey available (ref. 4 and refs therein). Therefore, the higher frequency of beak marks on large butterflies might indicate that they are more subject to predation⁵. However, my data show that, although males are significantly larger ($\bar{x}$ = 38.007 mm) than females ($\bar{x}$ = 36.020 mm) ($d = 12.9$, $P < 0.001$), beak marks are equally frequent in both sexes ($X^2_1 = 0.4$, $P > 0.5$) and independently so in both morphs (for *dorippus*, $X^2_1 = 0.3$, $P > 0.5$; for *aegyptius*, $X^2_1 = 0.6$, $P > 0.3$). This

Table 1 Summary of the data relating to beak-marks on *D. chrysippus* at Dar es Salaam, 1974–5

Month	x_1 Capture rate ($N h^{-1}$)	x_2 Mean forewing length (mm)	x_3 per cent* <i>aegyptius</i>	y per cent* beak marked	N
November '74	8.84	34.87	10.0	1.4	76
December	6.32	36.52	5.5	1.4	86
January '75	15.46	37.12	6.2	0.9	206
February	17.19	36.20	6.3	3.1	283
March	19.40	36.10	6.5	9.7	167
April	9.75	35.94	18.8	5.8	78
May	9.64	36.34	25.8	5.3	170
June	12.89	37.61	36.1	15.1	174
July	14.87	39.43	36.5	8.3	275
August	18.17	38.54	35.5	13.7	106
September	11.75	38.49	27.8	16.5	70

* arcsin transformed for statistical tests.

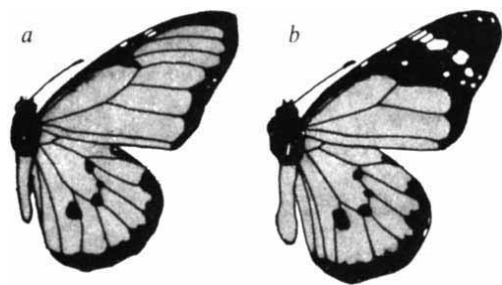


Fig. 1 *Danaus chrysippus* morphs *dorippus* (a) and *aegyptius* (b). Stippled areas are nut-brown to tawny orange, the remainder being black and white as shown.

unexpected result suggests that the same proportion of each sex is captured: moreover, females, though smaller and presumably weaker, escape predation after capture as readily as males.

The commonest food-plants (all Asclepiadaceae) available to *D. chrysippus* at Dar es Salaam⁶ are *Calotropis gigantea*, which contains cardiac glycosides (cardenolides)³ and *Tylophora stenoloba* which is non-toxic⁷. Butterflies (both sexes) reared on the former in the laboratory are significantly larger⁶. (Cardenolides are ingested by the caterpillar and stored in subsequent stages rendering them unpalatable to birds^{3,8-10}. In *D. plexippus*, cardenolides are especially concentrated in the wings¹¹ where they are likely to be tasted by the bird before the body is fatally damaged.) Thus, the similar proportion of beak marks in both sexes of *D. chrysippus* probably results from the larger individuals of each sex containing more cardenolide, being more distasteful and, consequently, more often rejected unharmed. There is no evidence that the larger size of males serves to protect females.

As the morphs do not differ significantly in size in either sex, they probably use the same range of food-plants⁶ and share a common palatability spectrum. When reared on *C. gigantea* in the laboratory at Dar es Salaam, the two morphs contained similar amounts of cardenolide³. Therefore, the higher frequency of beak marking on both male ($X^2_1 = 10.9$, $P < 0.001$) and female ($X^2_1 = 4.1$, $P < 0.05$) *aegyptius* is unlikely to result from palatability differences between the morphs. Moreover, within the range of my data, neither the density of the species as a whole nor the density or frequency of *aegyptius* itself had a significant effect on its beak-mark frequency.

The safety of a model is threatened if it has too many relatively edible mimics^{12,13}. At Dar es Salaam, the commonest mimic of *D. chrysippus* is *Hypolimnys misippus* L. (Nymphalidae) which has two female morphs, *misippus* and *inaria*, mimicking *aegyptius* and *dorippus* respectively¹⁴⁻¹⁵. If *H. misippus* is relatively poorly protected, the *misippus* form, which is much more common than *inaria*¹⁵, may constitute a threat to its *aegyptius* model by rendering the association as a whole relatively palatable. *H. misippus* females are rarely beak marked (3.2%, $n = 94$ in this study), suggesting they are usually consumed after capture. Furthermore, several common species of the well protected family Acraeidae¹⁶ closely resemble *dorippus* and are, therefore, its müllerian co-mimics. *Acraea encedon* alone has forms mimicking both *D. chrysippus* morphs. Thus, *dorippus* is probably better protected and, consequently, relatively avoided and less beak marked by virtue both of its more common müllerian and rare batesian mimics.

A more detailed analysis and discussion of these results will be published elsewhere.

Received 11 May; accepted 2 August 1979.

1. Swynnerton, C. F. M. *Jl S. Afr. Orn. Un.* **11**, 32-108 (1915).
2. Carpenter, G. D. H. *Proc. zool. Soc. Lond.* **111A**, 223-231 (1941).
3. Rothschild, M., van Eeuw, J., Reichstein, T., Smith, D. A. S. & Pierre, J. *Proc. R. Soc. B190*, 1-31 (1975).
4. Rothschild, M. in *Ecological Genetics and Evolution* (ed. Creed, R.) 202-223 (Blackwell, Oxford, 1971).
5. Edmunds, M. *Oikos* **25**, 117-118 (1974).
6. Smith, D. A. S. *Experientia* **34**, 845 (1978).
7. Abisch, E. & Reichstein, T. *Helv. chim. Acta* **45**, 2090-2116 (1962).

8. Brower, L. P., Edmunds, M. & Moffit, C. M. *J. Ent.* **A49**, 183-196 (1975).
9. Brower, L. P. *Scient. Am.* **220**(2), 22-29 (1969).
10. Brower, L. P., Brower, J. van Z. & Corvino, J. M. *Proc. natn. Acad. Sci. U.S.A.* **57**, 893-898 (1967).
11. Brower, L. P. & Glazier, S. C. *Science* **188**, 19-25 (1975).
12. Fisher, R. A. *The Genetical Theory of Natural Selection* (Oxford University Press, 1930).
13. Brower, L. P., Cook, L. M. & Croze, H. J. *Evolution* **21**, 11-23 (1967).
14. Smith, D. A. S. *Nature* **242**, 129-131 (1973).
15. Smith, D. A. S. *Biol. J. Linn. Soc.* **8**, 183-204 (1976).
16. Owen, D. F. *Oikos* **21**, 333 (1970).

Brassinolide, a plant growth-promoting steroid isolated from *Brassica napus* pollen

Michael D. Grove*, Gayland F. Spencer & William K. Rohwedder

Northern Regional Research Center, Science and Education Administration, US Department of Agriculture, Peoria, Illinois 61604

Nagabhushanam Mandava*, Joseph F. Worley, J. David Warthen Jr & George L. Steffens

Beltsville Agricultural Research Center, Science and Education Administration, US Department of Agriculture, Beltsville, Maryland 20705

Judith L. Flippen-Anderson

Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, District of Columbia 20375

J. Carter Cook Jr

School of Chemical Sciences, University of Illinois, Urbana, Illinois 61801

Extracts of *Brassica napus* L. (rape) pollen (family, Cruciferae) have been reported to produce a novel growth-promoting effect when applied to young pinto bean plants^{1,2}. Fractions that are active in the bean second-internode assay³ give rise to both increased cell elongation and cell division⁴. The validity of these early reports has been questioned because biological activity was not correlated with any one pure compound⁵. Subsequently, we showed that isolation of the growth promoter was complicated by the presence of chromatographically similar, inactive components that were identified as 6-D-glucopyranosyl esters of long-chain fatty acids⁶. We now report the isolation, structure determination and biological activity of this new plant growth promoter, named (22R, 23R, 24S)-2 α -3 α ,22,23-tetrahydroxy-24-methyl-6,7-s-5 α -cholestano-6,7-lactone or brassinolide (Fig. 1).

A partially refined 2-propanol extract⁷ of bee-collected rape pollen was subjected to successive chromatographic separations on silica gel (chloroform-methanol, 9:1 v/v; toluene-absolute ethanol, 9:1), on C₁₈ Hi-flosil (methanol-water, 6:4), and by high-performance liquid chromatography on μ -Bondapak C₁₈ (methanol-water, 65:35). The bean second-internode bioassay³ was used to guide fractionation. Brassinolide (4 mg from 40 kg of pollen) was crystallised (MeOH) and has the following physical properties: C₂₈H₄₈O₆ [high-resolution field desorption mass spectrometry shows m/e 481.3489 (M+1)⁺, calculated 481.3528]; melting point 274-275 °C; IR (KBr) 3,450, 1,690 cm⁻¹, (CH₂Cl₂) 1,720 cm⁻¹; ¹H-NMR (C₅D₅N) 80.72 (s, 3H), 1.03 (d, 6H, $J = 6.3$ Hz), 1.04 (s, 3H), 1.10 (d, 3H, $J = 6.4$ Hz), 1.13 (d, 3H, $J = 6.5$ Hz); ¹³C-NMR (CD₂Cl₂-CD₃OD, 9:1) 81.3, 11.9, 12.1, 15.6, 20.8, 21.0 (CH₃), 68.3, 68.4, 71.0, 73.7, 74.8 (C-O), 177.6 (C=O).

* To whom correspondence should be addressed.

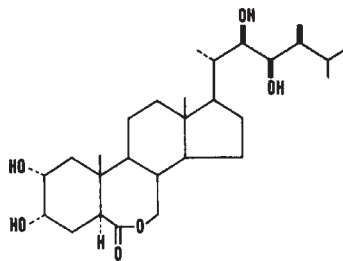


Fig. 1 Structure of brassinolide.

Field desorption mass spectrometry produced fragment ions at m/e 463 ($M+1-H_2O$)⁺, 445 ($M+1-2H_2O$)⁺, 409, 379, 361, 349, 131, 101 and 71. Chemical ionisation mass spectrometry (CI-MS) (isobutane reagent gas) showed ions at m/e 481, 463, 445, 427 ($M+1-3H_2O$)⁺, 409, 379, 361, 349, 343, 321, 303 and 279. High-resolution electron impact mass spectrometry showed ions for $C_{22}H_{36}O_5$ at m/e 380.2561 (calculated 380.2562) and $C_{22}H_{33}O_4$ at m/e 361.2398 (calculated 361.2378). Acetylation of brassinolide produced a tetraacetate: m/e 649 ($M+1$)⁺ by CI-MS. The mass spectral fragmentation pattern is consistent with cleavages along a dihydroxylated side chain and D-ring of a steroid (Fig. 2).

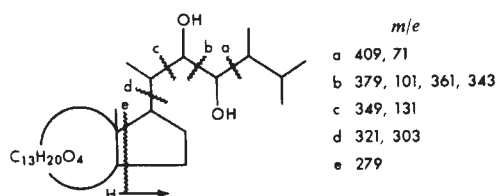


Fig. 2 Mass spectral fragmentation of brassinolide.

The structure of brassinolide was completely established by X-ray analysis of a single crystal which belonged to the monoclinic space group $P2_1$ with $a = 9.88(2)$ Å, $b = 7.63(2)$ Å, $c = 17.98(3)$ Å, and $\beta = 91.9(1)^\circ$. There is one molecule of brassinolide and one molecule of water in the asymmetric unit. A partial structure was obtained by application of the symbolic addition procedure for non-centrosymmetric crystals⁸. The fragment was developed into the full structure by the tangent formula refinement and expansion method⁹. A perspective drawing of brassinolide less hydrogens (Fig. 3) was obtained using experimentally determined atomic positions at an intermediate stage of least-squares refinement. Although only the relative configuration was determined by the X-ray analysis, the enantiomer chosen for Fig. 3 depicts the normal absolute configuration found in natural steroids.

When tested in the bean second-internode bioassay, crystalline brassinolide produced an approximately 200% increase in elongation of the internode over that of an untreated control

Table 1 Bean second-internode activity of brassinolide

Brassinolide (ng per plant)	Percentage increase in elongation over untreated control	No. of split internodes
1,000	235	3 of 3
100	189	3 of 3
10	182	1 of 3
1	21	0 of 3

Bioassay was carried out as previously described^{3,4}. Brassinolide was dissolved in dichloromethane-2-propanol before mixing with the fractionated lanolin carrier. Three plants were used for each concentration of brassinolide. An equal number of plants treated with lanolin carrier alone served as the control.

(Table 1). The increased cell division response was manifested by curvature, swelling and, most dramatically, by splitting of the second internode (Fig. 4) at concentrations of brassinolide as low as 10 ng per plant. Treatment with alkali destroys activity, which is regenerated on acidification.

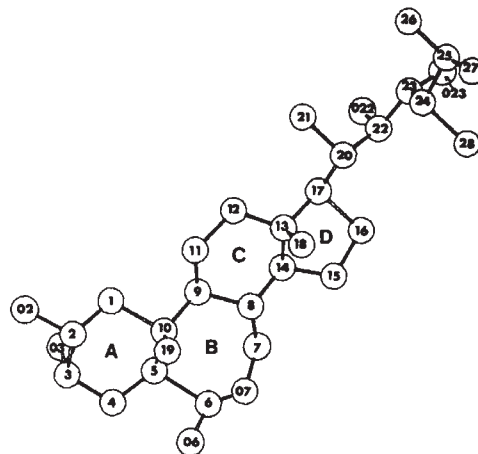


Fig. 3 Computer-generated perspective drawing of brassinolide. Hydrogens are omitted for clarity.

The brassinolide structure can be compared with ecdysone insect moulting hormones¹⁰. The 2,3-*cis*-glycol of brassinolide is α - rather than β -orientated; brassinolide has a *trans* A/B-ring fusion whereas ecdysones are *cis*-fused; ecdysones have a 7-ene-6-one grouping, whereas brassinolide has a B-ring lactone which is unprecedented in a natural steroid. Although the biosynthetic source of brassinolide is unknown, another C_{28} steroid, 24-methylencholesterol, comprises 56% of the rape pollen sterols¹¹. The 6,7-lactone functionality could originate from a 6-keto steroid precursor through a Baeyer-Villiger-type reaction. Oxidation of 3 β -acetoxycholestan-6-one was reported to result in migration of either C-7 or C-5, with the brassinolide-like 6,7-lactone predominating¹².

Fig. 4 Excised pinto bean (*Phaseolus vulgaris*) second internodes illustrating from left to right: untreated control; extremely swollen brassinolide treated; split brassinolide treated.

We thank D. W. Spaulding, M. V. Wakeman and B. C. Weathington for technical assistance, Dr D. Weisleder, W. R. Lusby and Dr H. Fineold for providing some of the spectral data.

Received 17 May; accepted 1 August 1979.

- Mitchell, J. W., Mandava, N., Worley, J. F., Plimmer, J. R. & Smith, M. V. *Nature* **225**, 1065-1066 (1970).
- Mitchell, J. W. & Gregory, L. E. *Nature new Biol.* **239**, 253-254 (1972).
- Mitchell, J. W. & Livingston, G. A. *Agric. Handb. U.S. Dept. Agric.* **336** (1968).
- Worley, J. F. & Mitchell, J. W. *J. Am. Soc. hort. Sci.* **96**, 270-273 (1971).
- Milborrow, B. V. & Pryce, R. J. *Nature* **243**, 46 (1973).
- Grove, M. D. *et al. Phytochemistry* **17**, 1187-1189 (1978).
- Mandava, N. *et al. Ind. Engng Chem., Prod. Res. Dev.* **17**, 351-354 (1978).
- Karle, J. & Karle, I. L. *Acta crystallogr.* **21**, 849-859 (1966).
- Karle, J. *Acta crystallogr.* **B24**, 182-186 (1968).
- Hikino, H. & Hikino, Y. *Fortschr. Chem. org. NatStoffe* **28**, 256-312 (1970).
- Knights, B. A. *Phytochemistry* **7**, 1707-1708 (1968).
- Cookson, R. C., Gandhi, R. P. & Southam, R. M. *J. chem. Soc. C*, 2494-2500 (1968).

Loss of immune competence with age may be due to a qualitative abnormality in lymphocyte membranes

Robin E. Callard* & Antony Basten

Immunology Unit, Department of Bacteriology, University of Sydney, NSW, 2006 Australia

Robert V. Blanden

John Curtin School of Medical Research, Australian National University, Canberra, 2600 Australia

The loss of immune function which occurs with age has until recently been attributed to depletion of immunocompetent cells¹⁻³. However, studies from our laboratory⁴⁻⁶ have indicated that the age-related decline in immune responsiveness is better explained by a qualitative rather than a quantitative defect at the level of both T cells and B cells. This conclusion was derived from experiments in which immune function was assessed in relation to the number of responding cells. Thus, lymphoid tissues from old and young mice were found to contain comparable numbers of T cells and B cells despite a 10-fold or greater difference in their antibody-forming potential and mitogenic activity. Furthermore, selective depletion of reactive subpopulations was excluded by the demonstration of similar proportions of mitogen and antigen-binding cells in lymphoid tissue from old and young mice. If the number of cells is normal in old animals, but their function is impaired, it is possible that the decline in immune competence with age may be due to a metabolic or structural abnormality, perhaps affecting the cell membrane and preventing the cell from responding normally to exogenous signals. We have therefore investigated whether any difference could be detected in membrane-associated antigens of lymphoid cells from old and young mice of the same strain. We report here that new or altered antigenic determinants seem to be expressed on the surface of lymphocytes from old mice. These might interfere with the normal immunological activities of such cells, and so perhaps explain their reduced immune competence.

Three separate techniques were used in our studies. First, spleen cells from six young (3-4 months), C57BL/6J mice were cultured with spleen cells from six syngeneic old (24-30 months) mice of the same sex in one-way mixed lymphocyte reactions (MLRs). Cells were prepared separately from each mouse and cultured in all possible responder-versus-stimulator combinations. As shown in Table 1, except for three young→old combinations, spleen cells from each of the six young donors responded to spleen cells from each of the six old mice. Conversely, cells from each old mouse responded to stimulator cells from each young donor. At no time was any response obtained in young→young and old→old combinations. These findings suggest that cells from old mice possess cell-surface determinants which can stimulate proliferation of young cells and vice versa. This observation was confirmed by the demonstration of a similar young→old MLR with spleen cells from another strain, CBA/Ca/T₆ (Table 2). Furthermore, the addition of polyethylene glycol, which potentiates weak MLRs⁷, significantly enhanced the young→old syngeneic MLR of cells from both C57BL/6J and CBA/Ca mice (Table 2). This supports the notion that the response obtained in young→old combinations of syngeneic cells represents a true MLR elicited by weak antigenic differences.

The second approach used for the detection of differences in membrane-associated antigens involved raising an antiserum to lymphoid cells from old mice in young syngeneic recipients (young anti-old). This young anti-old serum bound to 7-8% of spleen and lymph node cells from old mice but to only 1-2% of lymphoid cells from young mice (Table 3). Neither young anti-young control serum nor normal sera collected from both groups of recipients before immunisation bound to more than 2% of cells from old or young mice; this is no greater than the number of cells which label with protein A alone. It may be due to the presence of protein A antigen-binding cells or binding of protein A to membrane-associated IgG which is normally present on 1-2% of normal mouse spleen cells⁸. The ability of young anti-old serum to recognise determinants on old lymphoid cells was confirmed by complement-dependent cytotoxicity by the method of Boyse *et al.*⁹ (Table 3).

Table 1 Syngeneic one-way mixed lymphocyte reaction between spleen cells from old and young mice

Source of responder spleen cells	Source of stimulator spleen cells											
	Young mice						Old mice					
	1	2	3	4	5	6	1	2	3	4	5	6
Young mice												
1	—	0	0	0	0	0	1	2	5	1	0	4
2	0	—	0	0	0	0	1	3	4	1	3	3
3	<1	0	—	0	0	0	2	2	5	2	0	3
4	0	0	0	—	0	0	1	3	3	1	0	4
5	0	0	0	0	—	0	2	3	3	2	3	2
6	<1	<1	0	0	0	—	3	4	2	3	2	6
Old mice												
1	4	2	5	1	5	4	—	0	0	0	0	0
2	4	3	6	6	7	5	0	—	0	0	0	0
3	3	4	3	4	1	1	0	0	—	0	0	0
4	6	4	4	6	5	3	0	0	0	—	0	0
5	4	3	3	2	1	3	0	0	0	0	—	0
6	4	4	4	4	5	4	0	0	0	0	0	—

Spleen cells were prepared separately from six young and six old C57BL/6J mice of the same sex and cultured in all possible combinations in a one-way MLR. Quadruplicate cultures were carried out in Linbro microtitre trays (Flow Laboratories). Each well contained 5×10^5 stimulator cells and 5×10^5 responder cells in 0.2 ml medium RPMI 1640 (GIBCO) containing 2.5% fetal calf serum (FCS) (Flow Laboratories), 5×10^{-5} M 2-mercaptoethanol, 20 mM HEPES (Flow Laboratories), 2 g l^{-1} NaHCO₃, 60 µg ml^{-1} penicillin and 100 µg ml^{-1} streptomycin. Cell division was prevented in stimulator cells by exposure to 1,500 rads irradiation from a ⁶⁰Co source. Cultures were incubated at 37 °C in an humidified atmosphere of 5% CO₂ in air for 48 h, pulsed with 0.5 µCi of ³H-thymidine and collected 24 h later with a Skatron automatic harvester. Incorporation of ³H-thymidine was determined by counting in 0.5% diphenylloxazole/toluene scintillant using an LKB Wallac 81000 liquid scintillation counter. Counts were corrected for efficiency and quenching and expressed as d.p.m. The results are expressed in significant ($P < 0.05$) d.p.m. $\times 10^{-3}$ above background (—). Background counts were obtained from cultures of spleen cells with irradiated cells from the same animal and were generally less than 5,000 d.p.m. Cultures in which d.p.m. were not significantly above background are expressed as 0. It can be seen that in all but three combinations, cells from each young mouse (young 1-6) responded significantly above background to cells from each of the six old donors. Similarly, cells from each of the six old donors (old 1-6) responded to young but not old cells.

The third group of experiments was designed to test whether the new determinants expressed on old cells might involve antigens of the major histocompatibility complex (MHC), in view of their known importance in modulating immune responsiveness. Reciprocal skin grafts were therefore exchanged between old and young C57BL/6J mice of the same sex. These, however, remained intact for longer than 60 d, suggesting that any modification of cell-surface components with age did not result in significant alteration to either major or minor histocompatibility antigens involved in graft rejection. Possible changes in H-2K and H-2D antigens were further tested for by using an 'H-2-restricted' cytotoxic T-cell response to viral infection¹⁰. Cytotoxic T cells specific for H-2K or H-2D plus viral antigens were generated in young C57BL/6J mice. These cells lysed virus-infected peritoneal macrophage target cells from old and young mice with similar efficiency (Table 4). As this assay is very sensitive to qualitative changes in H-2K or H-2D antigens caused by mutations^{11,12}, the negative results obtained

* Present address: ICRF Human Tumour Immunology Group, University College Hospital Medical School, University Street, London WC1, UK.

Table 2 Enhancement of syngeneic young → old MLR with polyethylene glycol

Source of spleen cells used in syngeneic MLR	Polyethylene glycol added	
	No	Yes
C57BL/6		
Young → old	2,700 ± 484	8,100 ± 1,750
CBA		
Young → old	9,900 ± 1,300	14,370 ± 3,900

Syngeneic young → old MLRs using spleen cells from two strains of mice were carried out with and without 5% polyethylene glycol, which is known to enhance weak MLRs. Results are expressed as d.p.m. above background counts obtained from young → young MLR cultures. Young mice were aged 3–6 months and old mice 24–30 months.

suggest that an age-related change in H-2K or H-2D antigens is unlikely, but do not exclude alterations in I-region antigens which could explain the stimulation seen in the MLR experiments (Tables 1, 2).

The above findings suggest that new or altered antigenic determinants are expressed on the surface of lymphoid cells from old mice. Antigenic differences between populations of syngeneic lymphocytes have also been reported previously in other systems. For example, cell-mediated cytotoxicity has been obtained in female mice against male-specific (H-Y) antigens¹³, although such an explanation has been excluded in the present study by the exclusive use of mice of one sex. In addition, mixed lymphocyte reactivity has been demonstrated with combinations of neonatal thymocytes and syngeneic adult spleen cells¹⁴, which is consistent with the appearance of new antigens during cellular differentiation. More recently, Gozes *et al.*¹⁵ have demonstrated a syngeneic graft-versus-host (GvH) response induced in popliteal lymph nodes of young C57BL/6 mice by footpad injection of syngeneic spleen cells from old donors. This finding, as well as the young → old syngeneic responses reported

Table 3 Specific binding and complement-mediated cytotoxicity of young anti-old serum against lymphoid cells from old mice but not young mice

Donor mice	Source of cells	% Of lymphocytes labelled with:		
		Young anti-old serum	Young anti-young serum	Normal mouse serum
Old	Spleen	7.5 ± 1.2*	2	2
Old	Lymph node	6.8 ± 2.0	2	2
Young	Spleen	2	2	2
Young	Lymph node	2	2	2
		% Of lymphocytes killed with:		
		Young anti-old serum	Young anti-young serum	Normal mouse serum
Old	Spleen	14.2 ± 1.2	7.0 ± 0.6	6.0 ± 0.6
Young	Spleen	7.0 ± 1.3	6.4 ± 1.1	6.0 ± 0.8

Antisera specific for old lymphoid cells were raised in syngeneic young recipients of the same sex by six injections of 10⁷ pooled spleen and lymph node cells from old C57BL/6J donors given intraperitoneally at weekly intervals. The first, third and fifth injections were given as an emulsion in Freund's complete adjuvant, and viable cells were used on the remaining occasions. A control serum (young anti-young) was prepared in the same way except that young spleen and lymph node cells were substituted for old cells in the immunisation regime. Both antisera were collected 7 d after the final injection. All sera were tested on lymphoid cells from syngeneic mice of the same sex. Specific binding of anti-lymphocyte antibodies was determined autoradiographically, using ¹²⁵I-labelled protein A, which binds to the Fc portion of immunoglobulin²³. Briefly, 5 × 10⁶ spleen or lymph node cells were incubated with 20 µl of antiserum in 100 µl of medium 199 containing 0.5% bovine serum albumin (Commonwealth Serum Laboratories) and 15 mM sodium azide for 30 min on ice. After two washes, the cells were incubated with 0.5 µg of ¹²⁵I-protein A (protein A obtained from Pharmacia and iodinated by the method of Greenwood *et al.*²³ to a specific activity of 10 µCi per µg) in 0.4 ml medium. Unbound label was removed by centrifugation and the cells smeared on gelatin-coated slides. Slides were dipped in Kodak NTB2 emulsion and exposed for 10–21 d at 4 °C and then developed in Kodak D19. The cells were stained with Giemsa and the proportion of labelled small mononuclear cells determined from counts of 1,000 cells per slide. Complement-dependent cytotoxicity was carried out according to the method of Boyse *et al.*⁹.

* Mean value ± s.e.m.

here, could represent an autoimmune response to 'self' antigens or a response to new antigenic determinants on old lymphoid cells. Such new antigenic determinants could have arisen in various ways; for example, by changes in the cholesterol/phospholipid ratio in the plasmalemma (of the type reported by Rivnay *et al.*¹⁶) which alters membrane viscosity, disturbs normal lymphocyte function and may expose hidden antigenic determinants, or by insertion, into the lymphoid cell membrane, of altered proteins arising by somatic mutation or from frequent errors in protein synthesis, both of which have been postulated as causes of cellular ageing^{17,18}.

Table 4 Lysis of ectromelia virus-infected peritoneal macrophages from old or young C57BL/6J mice by immune spleen cells from young syngeneic donors

Source of target cells	⁵¹ Cr Release		
	Water lysis	Spontaneous lysis	Lysis caused by immune spleen cells
Virus-infected macrophages from old mice	93.5 ± 0.5	34.5 ± 2.0	72.4 ± 2.1* 51.7 ± 3.7*
Control (uninfected) macrophages from old mice	90.1 ± 1.1	22.8 ± 1.2	21.9 ± 0.1 27.0 ± 0.8
Virus-infected macrophages from young mice	92.8 ± 1.4	33.6 ± 2.3	79.3 ± 0.8 46.4 ± 2.2
Control (uninfected) macrophages from young mice	87.4 ± 2.2	24.0 ± 0.8	24.4 ± 1.5 24.7 ± 1.0

* Results obtained from two individual immune mice given 10⁴ plaque-forming levels of ectromelia virus intravenously 5 d previously. The killer: target ratio was 20:1 and the assay was run for 16 h.

† Macrophages were infected with 40 plaque-forming units of virus for 1 h at 37 °C and left for 3 h before addition of immune spleen cells as described in detail elsewhere¹¹.

The nature of the new antigenic determinants detected on lymphoid cells from old mice is unknown but the current experiments do shed light on some of the possibilities. The lack of skin graft rejection between old and young mice and the failure to detect differences in cytotoxic T-cell activity (Table 4) argue against alterations in K or D antigens of the MHC. However, neither experiment excludes changes in Ia or M locus antigens which may not give rise to graft rejection. Consistent with modifications in I-region determinants is the finding of mixed lymphocyte reactivity between old and young cells from two different strains of mice (Tables 1, 2). Moreover, further evidence suggests altered expression of Ia antigens in old mice. For example, macrophages from old mice are abnormal in their ability to function as stimulator cells in an allogeneic MLR¹⁹ and an increase in I-J-bearing T cells corresponding to the presence of T-cell suppressors has been detected in old mice (R.E.C. *et al.*, in preparation). An alteration in I-region products with age could explain the decline in lymphocyte function with age, as Ia antigens have a key role in cellular interactions of the immune response. Also, modified Ia antigens may contribute to the high incidence of neoplasia and autoimmunity in the aged. This is supported by the finding of Gleichmann *et al.*²⁰ of an increased incidence of lymphomas and autoimmunity in chronic GvH disease involving I-region differences only. In addition, the rise in I-J-bearing suppressor cells may further exacerbate the development of tumours²¹. These membrane changes may also explain the high incidence of autoimmune disease in the aged: insertion of a new, non-self, antigen onto the lymphocyte cell surface could allow T-cell help to be expressed for antibody responses to other, self surface antigens. The alternative explanation of autoimmune reactivity to self antigens resulting from a loss of T-cell suppressor function seems less likely

because an increase rather than a decrease in T-cell suppressor activity occurs in old mice^{6,22}. Whatever the exact nature of the new antigen determinant on old lymphoid cells, it is clear that it may profoundly affect normal immune functions. Furthermore, accumulation of qualitative abnormalities of this kind in other somatic cells may account for ageing in a wide variety of higher organisms.

This work was supported by grants from the National Health and MRC of Australia and the University of Sydney Cancer Research Funds. We thank Dr I. F. C. McKenzie for doing the skin grafts.

Received 3 May; accepted 8 August 1979.

1. Burnet, F. M. *Lancet* **ii**, 358 (1970).
2. Albright, J. F. & Makinodan, T. *J. cell. Physiol.* **67**, Suppl. 1, 185 (1966).
3. Kruisbeek, A. M. *Mechanisms Ageing Dev.* **5**, 125 (1976).
4. Callard, R. E. & Basten, A. *Cell. Immun.* **31**, 13 (1977).
5. Callard, R. E., Basten, A. & Waters, L. K. *Cell. Immun.* **31**, 26 (1977).
6. Callard, R. E. & Basten, A. *Eur. J. Immun.* **8**, 552 (1978).
7. Ben-Sasson, S. A. & Henkart, P. A. *J. Immun.* **119**, 227 (1977).
8. Uhr, J. W. in *Membrane Receptors of Lymphocytes* (eds Seligmann, M., Preud'homme, J. L. & Kourilsky, F. M.) 3–12 (Masson, Paris, 1975).
9. Boyse, E. A., Old, L. J. & Chouroulinkou, I. *Meth. med. Res.* **10**, 39 (1964).
10. Doherty, P. C., Blanden, R. V. & Zinkernagel, R. M. *Transplant Rev.* **26**, 89 (1976).
11. Blanden, R. V., Dunlop, M. B. C., Doherty, P. C. & Kohn, H. I. *Immunogenetics* **3**, 541 (1976).
12. Zinkernagel, R. M. & Klein, J. *Immunogenetics* **4**, 581 (1977).
13. Gordon, R. D., Simpson, E. & Samelson, L. E. *J. exp. Med.* **142**, 1108 (1975).
14. von Boehmer, H. & Adams, P. J. *Immun.* **110**, 376 (1973).
15. Gozes, Y., Umiel, T., Meshorer, A. & Trainin, N. *J. Immun.* **121**, 2199 (1978).
16. Rivnay, B., Globerson, A. & Shinitzky, M. *Mechanisms Ageing Dev.* **10**, 71 (1979).
17. Burnet, F. M. *Intrinsic Mutagenesis: A Genetic Approach to Ageing* (Medical and Technical, St Leonards, 1974).
18. Orgel, L. E. *Nature* **243**, 441 (1973).
19. Callard, R. E. *Eur. J. Immun.* **8**, 698 (1978).
20. Gleichmann, E., Gleichmann, H. & Wilke, W. *Transplant Rev.* **31**, 156 (1976).
21. Perry, L. L., Benacerraf, B. & Greene, M. I. *J. Immun.* **121**, 2144 (1978).
22. Callard, R. E., Fazekas De St. Groth, B., Basten, A. & McKenzie, I. F. C. *J. Immun.* (submitted).
23. Forsgren, A. & Sjoquist, J. *J. Immun.* **97**, 822 (1966).
24. Greenwood, F. C., Hunter, W. M. & Glover, J. S. *Biochem. J.* **89**, 114 (1963).

Immunity to human haematopoietic stem cells

Robert Peter Gale & Caprice Rutkosky Moran

UCLA Bone Marrow Transplant Team, UCLA School of Medicine, Center for the Health Sciences, Los Angeles, California 90024

Graft rejection is an important problem in human bone marrow transplantation, with a 25–60% incidence in grafts between HLA-identical siblings^{1–6}. In mice, two models of rejection of bone marrow grafts have been described: T-lymphocyte-mediated immunity and hybrid resistance (for reviews see refs 7, 8). The first model involves the recognition of foreign histocompatibility (H-2) antigens by selected subsets of T lymphocytes. These antigen-reactive cells undergo clonal expansion and differentiate into cytotoxic effector cells which mediate graft rejection. In the hybrid resistance model, mononuclear cells, probably identical to natural killer (NK) cells, react against haematopoietic histocompatibility antigens on parental marrow cells⁹. Cells involved in hybrid resistance exhibit 'spontaneous' cytotoxicity to disparate haematopoietic cells and, in contrast to T-lymphocyte-mediated immunity, do not require a sensitisation phase before differentiating into cytotoxic effector cells. The mechanism of graft rejection following bone marrow transplantation in man is unknown. In transplants between HLA-identical siblings the target antigens of graft rejection are not HLA or antigens closely linked to the major histocompatibility complex (MHC). To study the problem of graft rejection in man we developed and describe here an *in vitro* model to detect immunity to haematopoietic stem cells committed to granulocyte-macrophage differentiation (CFU-C; colony-

forming unit culture). We found that human peripheral blood and bone marrow but not thymus contain a population of cells which are cytotoxic to haematopoietic stem cells following *in vitro* sensitisation. The effector cell population is not a conventional T or B lymphocyte, monocyte-macrophage or NK cell.

The model system is shown in Fig. 1. Peripheral blood mononuclear cells (PBMC) were co-cultivated with irradiated bone marrow cells for 5–7 days. Sensitised PBMC were recovered and incubated with fresh, unirradiated bone marrow cells for an additional 16 h after which the PBMC-bone marrow mixture was suspended in semi-solid agar and plated into Petri dishes over an underlayer containing human colony-stimulating activity (CSA). Granulocyte-macrophage colonies (CFU-C) were counted 7–14 days later using an inverted microscope^{10,11}.

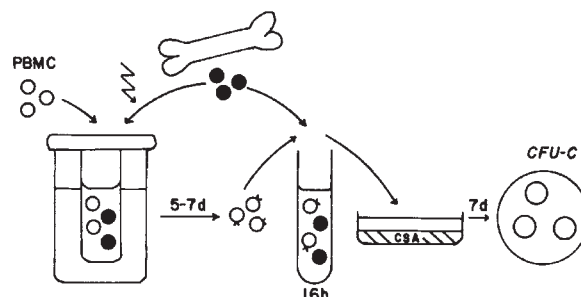


Fig. 1 Experimental model of immunity to granulocyte-macrophage progenitors (CFU-C). Peripheral blood mononuclear cells (PBMC), obtained by Ficoll-Hypaque density centrifugation of venous blood, were resuspended in McCoy's 5A medium, 20% fetal calf serum (FCS) and 1% penicillin-streptomycin (PS). Bone marrow from normal donors was obtained by needle aspiration from the posterior iliac crest into syringes containing preservative-free heparin. A leukocyte-rich bone marrow cell population was prepared by centrifugation in Wintrobe tubes at 1,600g for 10 min. 10^6 PBMC (0.5 ml) were co-cultivated with an equal number of irradiated bone marrow cells (1,500 rads, Gammacell-220 ^{60}Co -source; rate, 3,200 rad min^{-1}) in Marbrook flasks for 5–7 days at 37 °C in a 7.5% CO_2 atmosphere. Sensitised PBMC were recovered, washed and resuspended in McCoy's 5A, 20% FCS and 1% PS. Sensitised PBMC (10^6 , 1 ml) were cultured with an equal number of bone marrow cells in 16×125 mm plastic tubes for 16–24 h at 37 °C, 7.5% CO_2 atmosphere. Cells (2×10^5) were then diluted in 0.3% agar in McCoy's 5A with 15% FCS and plated over white blood cell underlayers in 0.5% agar in Petri dishes which were incubated for 10 d at 37 °C. Colonies of ≥ 40 cells were counted using an inverted microscope^{10,11}.

Data were expressed as the $\bar{X} \pm \text{s.e.m.}$ of triplicate cultures and were analysed by paired *t*-test or corrected χ^2 analysis as indicated. Representative data of five experiments are shown in Fig. 2. Normal human marrow has a cloning efficiency of 104 ± 9 CFU-C and normal PBMC less than 10 CFU-C per 2×10^5 cells plated in these conditions. The addition of PBMC co-cultured with autologous bone marrow to fresh, unirradiated bone marrow from unrelated donors consistently increased the number of CFU-C, probably related to CSA release. In contrast, PBMC sensitised to allogeneic bone marrow consistently inhibited the CFU-C in fresh bone marrow from donors identical or unrelated to the donor of the sensitising cells (Fig. 2). No CFU-C inhibition was observed when bone marrow from the PBMC donor was used as a target or in cultures between genetically identical twins. These data exclude a nonspecific effect of inhibition of CFU-C by *in vitro* sensitisation.

The interaction between sensitised PBMC and target CFU-C is time and dose dependent and requires cell-cell contact (Fig. 3). Twenty per cent inhibition of CFU-C was observed after 2–4 h and 80% after 16 h of secondary culture. A 2:1 ratio of PBMC to bone marrow usually resulted in a 70–80% inhibition of CFU-C. CFU-C inhibition was not observed when sensitised PBMC were separated from target bone marrow cells by either a Nucleopore or Millipore membrane (control 104 ± 6 ; co-culture 22 ± 3 ; separated culture 97 ± 7).

Cell-mediated inhibition of CFU-C was relatively unaffected by depletion of T lymphocytes by anti-thymocyte globulin

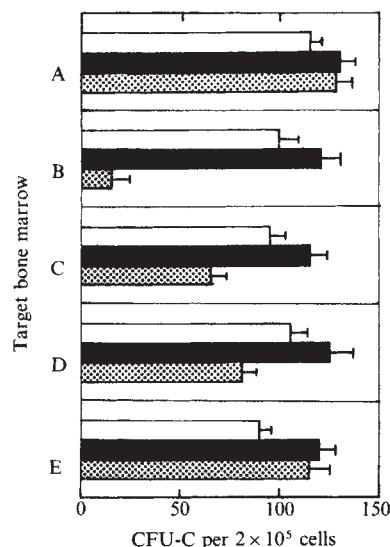


Fig. 2 CFU-C inhibition by sensitised effector cells. PBMC from individual A were sensitised to autologous bone marrow cells (black bars) or to bone marrow from an unrelated individual B (stippled bars). The sensitised cells were then co-cultured with fresh bone marrow cells and CFU-C determined. Target bone marrow from A and B and from three unrelated individuals (C, D, E) were also tested (ordinate). CFU-C content of the bone marrow from each donor without addition of sensitised effector cells is indicated by the clear bars.

(ATG), or removal of sheep erythrocyte rosette-forming cells (E-RFC); by removal of B lymphocytes by nylon wool columns or with anti-B-lymphocyte antisera; or by removal of monocyte-macrophages by phagocytosis of carbonyl iron or adherence to plastic surfaces (Fig. 4). Concentrations of silica and carrageenan that completely inhibited NK activity also had no effect on PBMC-mediated CFU-C inhibition. Furthermore, purified populations of T and B lymphocytes, monocyte-macrophages and NK cells showed no inhibition of CFU-C in identical conditions (Table 1). Bone marrow-derived mononuclear cells but not thymocytes could also serve as a source of effector cells in this assay (data not shown). These data indicate that the effector cell is probably not a conventional T or B lymphocyte, monocyte-macrophage or NK cell.

The target antigen(s) on CFU-C recognised by sensitised PBMC is unknown. To evaluate the role of HLA antigens, we studied five pairs of HLA-identical siblings. Figure 5 shows representative data. CFU-C inhibition was reproducibly demonstrated in four of the five pairs of siblings, indicating that HLA antigens are not the primary target of CFU-C inhibition. In most instances, the degree of CFU-C inhibitions between HLA-identical siblings was not as great as that observed between HLA-mismatched combinations. Recently, B-lymphocyte-associated antigens (HLA-DR or Ia-like) have

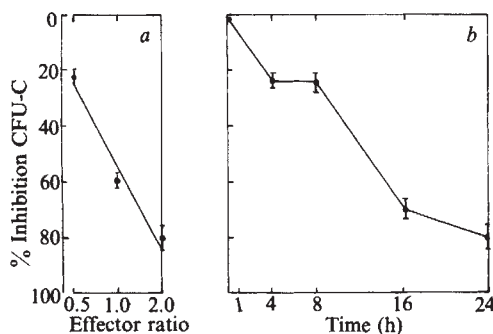


Fig. 3 Kinetics of CFU-C inhibition. Sensitised PBMC recovered from primary cultures were recultured with fresh bone marrow cells at various ratios (a) and for various lengths of time (b). The secondary cultures were then terminated and CFU-C assayed as indicated (Fig. 1).

Table 1 Effects of mononuclear cell subpopulation on CFU-C

	Mononuclear cell sub-population			
	T	B	M	NK
Control	86 ± 4	112 ± 7	96 ± 7	75 ± 4
Sensitised PBL	30 ± 2	32 ± 3	20 ± 3	17 ± 2
Depleted	32 ± 5	26 ± 3	17 ± 5	33 ± 6
Enriched	93 ± 6	102 ± 5	120 ± 10	102 ± 7

The CFU-C assay was carried out as described in Fig. 1 legend. PBMC from HLA-mismatched individuals were co-cultured and tested for inhibition of bone marrow cells from the same donor as the sensitising PBMC. Results are expressed as follows: control, autologous PBL co-cultures with target bone marrow cells ($A \times A \rightarrow B$); sensitised, allogeneic PBMC co-cultures with target BM ($A \times B \rightarrow B$). HLA-identical siblings were used in the experiment involving T lymphocytes. In other experiments HLA-mismatched, unrelated donors were used. Sensitised PBMC were fractionated using techniques described in Fig. 4 legend. Subpopulations were depleted or enriched for T lymphocytes (E-rosettes), B lymphocytes (nylon-wool column) or macrophages (adherence). NK cells were isolated by fractionation on an immune complex monolayer as described elsewhere²². Results are expressed as CFU-C per 2×10^5 cells plated.

been demonstrated on CFU-C^{12,13}. The possibility that this antigen was the target of CFU-C inhibition was investigated by experiments in which an antiserum to human B-lymphocyte antisera failed to block CFU-C inhibition by sensitised PBMC¹⁴ (control 114 ± 6 ; sensitised culture 27 ± 7 ; sensitised culture with anti-Ia, 21 ± 3). The CFU-C inhibition observed between HLA-identical siblings is further indirect evidence that B-lymphocyte-associated antigens are not important in this

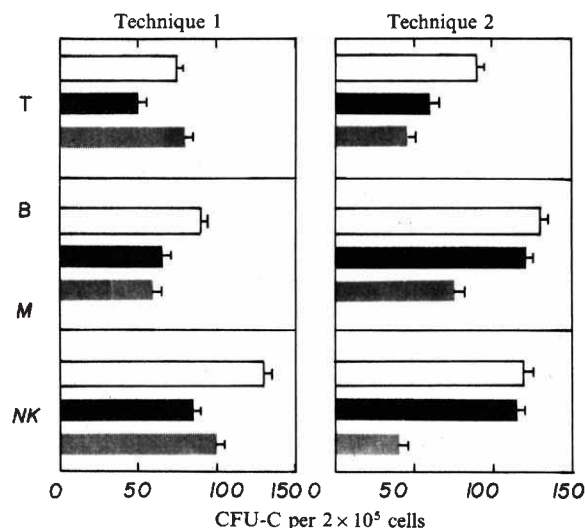


Fig. 4 Effects of cell fractionation on CFU-C inhibition in control cultures (stippled bars), sensitised cultures (black bars) and sensitised cultures depleted of selected mononuclear cell populations (striped bars) by the method indicated. Sensitised PBMC were fractionated using the procedures indicated. After each procedure the cells were washed, reconstituted to their original concentration and tested for CFU-C inhibition. Unfractionated sensitised PBMC served as a control. T-lymphocyte-depleted fractions were obtained by incubation of 10^7 PBMC (1.0 ml) with 0.25 ml equine anti-human thymocyte globulin (ATG; Upjohn-17912; $1:10^3$ dilution) and rabbit complement or by rosette formation with AET (2-aminoethylisothiourea bromide hydrobromide) treated sheep erythrocytes (E-RFC) and centrifugation over a Ficoll-Hypaque density gradient¹⁷. T-depleted PBMC contained $\leq 5\%$ E-rosette forming cells. B lymphocytes were depleted by adherence to nylon wool columns or by incubation with a rabbit anti-human B-lymphocyte antiserum and complement^{18,19}. B-lymphocyte-depleted fractions contained $\leq 2\%$ surface membrane Ig-positive (SMIg⁺) cells. Monocyte-macrophage-depleted (M) fractions were obtained by incubation with carbonyl iron (lymphocyte separating reagent, Technicon) and passage through a laboratory magnet or by adherence to polystyrene surfaces. Macrophage-depleted fractions contained $\leq 5\%$ nonspecific esterase staining cells. NK cells were inactivated by incubation with silica or carrageenan. PBMC were incubated for 30 min with silica, $0.1-1.5 \text{ mg ml}^{-1}$ (Sigma; particle size $<0.4 \mu\text{m}$). The silica was sonicated and used uncoated or after incubation with 10% FCS. PBMC were also incubated with carrageenan, 0.15 mg ml^{-1} (Hathaway Allied Products) for 60 min. Treated fractions were tested for NK activity against sensitive (K-562, MOLT-4) and resistant (RAJI) cell lines. In each instance NK activity was absent^{20,21}. HLA-identical siblings were used in the experiments involving T lymphocytes. In other experiments HLA-mismatched, unrelated donors were used.

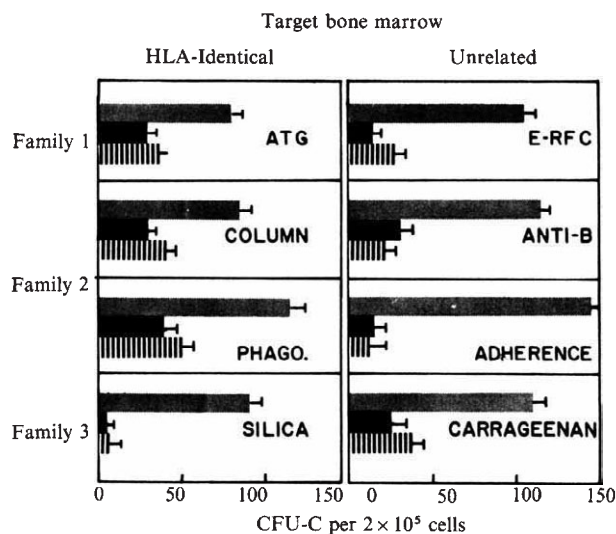


Fig. 5 CFU-C inhibition between HLA identical siblings. In three families with two HLA-identical siblings, PBMC from one sibling were sensitised against bone marrow cells from either the HLA-identical sibling (black bars) or an unrelated individual (stippled bars). Their sensitised PBMC were then tested for CFU-C inhibition by co-culture with fresh bone marrow from either the sibling (left panel) or the unrelated individual (right panel). The values indicated by the clear bars represents the number of CFU-C present when target bone marrow was cultured with PBMC sensitised against autologous bone marrow cells in primary culture.

system, as these antigens are closely linked or identical to the HLA-D locus shared by these siblings¹⁵.

We have thus demonstrated that PBMC sensitised against allogeneic but not autologous bone marrow contain a population of cells which inhibit haematopoietic stem cells committed to granulocyte-macrophage differentiation (CFU-C). The effector cell is not a conventional T or B lymphocyte, monocyte-macrophage or NK cell. The target antigen(s) on the CFU-C is unknown but is not HLA or a B-lymphocyte-associated (HLA-DR or Ia-like) antigen. This system should prove useful for studying immunity to non-HLA antigens and as a model for bone marrow graft rejection in man.

Dr R. Billings provided the anti-human B-lymphocyte antisera, the ATG was from Dr A. Elberg (UpJohn), Dr M. Tagasaki assayed NK activity, Dr P. Terasaki carried out the HLA typing, Drs E. Arenson, A. Saxon and S. Targan collaborated in the cell fractionation experiments, and Drs M. Cline and D. Golde critically reviewed the manuscript. This work was supported by grants CA-23175, CA-12800 and CA-15688 from the NCI and RR-865 from the USPHS. R.P.G. is a Scholar of the Leukemia Society of America.

Received 11 June; accepted 1 August 1979.

1. UCLA Bone Marrow Transplant Team *Lancet* **ii**, 921 (1976).
2. Storb, R., et al. *Transplant Proc.* **10**, 135 (1978).
3. Gluckman, F. et al. *Transplant Proc.* **10**, 141 (1978).
4. Elfenbein, G. H., Anderson, P. N., Klein, D. L., Schacter, D. Z. & Santos, G. W. *Transplant Proc.* **10**, 441 (1978).
5. ACS/NIH Bone Marrow Transplant Registry *J. Am. med. Ass.* **236**, 1131 (1976).
6. Camitta, B. M. et al. *Blood* **48**, 63 (1976).
7. Cerrottini, J. C. & Brunner, K. T. *Adv. Immun.* **18**, 67 (1974).
8. Cudkowicz, G., Landy, M. & Shearer, G. M., (eds) *Natural Resistance Systems Against Foreign Cells, Tumors, and Microbes* (Academic, New York, 1978).
9. Kiessling, R. et al. *Eur. J. Immun.* **7**, 655 (1977).
10. Golde, D. W. & Cline, M. J. *Blood* **41**, 45 (1973).
11. Pike, B. L. & Robinson, W. A. *J. cell. Physiol.* **76**, 77 (1970).
12. Cline, M. J. & Billings, R. J. *Exp. Med.* **146**, 1143 (1977).
13. Winchester, R. J. et al. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4102 (1977).
14. Billings, R. J., Honig, R., Peterson, P., Safani, M. & Terasaki, P. *J. Immun.* **117**, 2040 (1976).
15. Mann, D. L., Abelson, L., Henkart, P., Harris, S. D. & Amos, D. B. *Proc. natn. Acad. Sci. U.S.A.* **72**, 5103 (1975).
16. Golde, D. W. & Cline, M. J. *J. clin. Invest.* **51**, 2981 (1972).
17. Saxon, A., Feldhaus, J. & Robins, R. A. *J. immun. Meth.* **12**, 285 (1976).
18. Billings, R., et al. *J. exp. Med.* **144**, 167 (1976).
19. Billings, R., Ting, A. & Terasaki, P. *J. natn. Cancer Inst.* **58**, 199 (1977).
20. Takasugi, M., Mickey, M. R. & Terasaki, P. I. *Cancer Res.* **33**, 2898 (1973).
21. Takasugi, M., Mickey, M. R. & Terasaki, P. I. *J. natn. Cancer Inst.* **53**, 1527 (1974).
22. Targan, S. & Jondal, M. *J. immun. Meth.* **22**, 123 (1978).

H-2 restriction of contact inhibition of epithelial cells

A. S. G. Curtis & Paul Rooney

Department of Cell Biology, University of Glasgow, Glasgow, UK

Medawar has suggested¹ that the major histocompatibility gene complex (MHC) might be involved in the contact inhibition of movement² shown by fibroblasts and epithelia. Contact inhibition of movement³ is that reaction of cells which stops the movement of one cell over another and thus leads to monolayering and other features of morphology typical of cells in culture and *in vivo*⁴. By confronting epithelial outgrowths we have compared contact inhibition between syngeneic cells with that between allogeneic cells. Contact inhibition was more marked between allogeneic combinations than syngeneic combinations, when the genetic mismatch lay in certain parts of the MHC complex.

A number of cell interaction gene loci (CI loci) have been recognised in lymphocyte-lymphocyte and lymphocyte-macrophage interactions⁵, controlling such reactions as T-B lymphocyte cooperation in antibody responses, and the T-T lymphocyte reactions in the genesis of cytotoxic lymphocytes (CTL)⁶, as well as the reaction of the CTL with their target cell (which may be fibroblasts). These CI loci map in the MHC and, in mice, afford examples of H-2 restriction. If the CI loci actually control the interaction part of the process, rather than the response to the interaction, a question that has not been answered, we would already possess clear evidence that some cell interaction phenomena (rather than a consequent response) are under MHC control. There is evidence to show that cell adhesion of mouse fibroblasts is under H-2 control⁷, that the adhesion of mouse lymphoid cells is under the control of one or more of the loci K, IA, IB, D or L (though minor effects of others could not be excluded)⁸, that the adhesion of mouse T lymphocytes is controlled by loci D and/or L and perhaps IB, IC and IJ (ref. 9), and that mouse B lymphocyte adhesion is controlled by loci K and perhaps IB and IC (ref. 10). In refs 9 and 10, clear evidence was presented that allogeneic combinations of cells, mismatched at these loci, led to a diminution of adhesion, effected by soluble glycoproteins secreted by the cells.

Epithelia, which would be expected to show marked contact inhibition of movement¹¹, were grown as outgrowths of explants of adult mouse kidney tubule and the extent of contact inhibition measured (comparing the number of nuclear overlaps with the number expected on random grounds) in the area of confrontation between syngeneic or allogeneic outgrowths. The congenic strains B10, B10.A, B10.A(2R), B10.A(4R), B10.AKM, B10.G, B10.BR and B10.HTT were used, and also strain ATL. The results are shown in Table 1.

The results show that although there is a low overlap index (marked contact inhibition of movement) in the syngeneic controls, the allogeneic combinations show a very low overlap index, significantly below that of the controls, when certain loci are mismatched between the two strain types in the culture. In other words, there is an H-2 restriction affecting the extent of contact inhibition. If the cells involved are mismatched at K or/and D, contact inhibition is markedly increased.

It might be expected that in a culture dish containing equal numbers of explants from two strain types, 50% of the confronted outgrowths would be syngeneic and 50% allogeneic, and thus, that a population of overlap indices with two (possibly three) different means corresponding to two syngeneic and one allogeneic groups. Yet, all confronted outgrowths in such cultures show a very low overlap index (see Table 1.). A simple

Table 1 Contact inhibition of movement of mouse kidney tubule epithelium in allogeneic and syngeneic confrontations

Strain	Overlap index*	Haplotype										H-2 mismatch	t-test
Multi-explant cultures		K	IA	IB	IJ	IE	IC	S	G	D			
B10.G	0.27 (0.12)	q	q	q	q	q	q	q	q	q			
B10.BR	0.31 (0.24)	k	k	k	k	k	k	k	k	k			
B10.G+B10.BR	0.11 (0.07)										All loci	= 6.01	] = 4.55
B10.G	0.28 (0.12)	q	q	q	q	q	q	q	q	q			
B10.AKM	0.23 (0.13)	k	k	k	k	k	k	k	k	q			] = 7.47
B10.G+B10.AKM	0.08 (0.06)										All but D		
B10.BR	0.30 (0.12)	k	k	k	k	k	k	k	k	k			] = 4.69
B10.AKM	0.23 (0.13)	k	k	k	k	k	k	k	k	q			
B10.BR+B10.AKM	0.10 (0.06)										D only		] = 5.48
B10.A(2R)	0.31 (0.06)	k	k	k	k	k	d	d	d	b			
B10.A(4R)	0.34 (0.11)	k	k	b	b	b	b	b	b	b			] = 0.72
B10.A(2R)+B10.A(4R)	0.34 (0.19)										IB, IJ, IE, IC, S & G		
B10.HTT	0.29 (0.09)	s	s	s	s	k	k	k	k	d			] = 0.10
ATL	0.36 (0.12)	s	k	k	k	k	k	k	k	d			
B10.HTT+ATL	0.30 (0.12)										IA, IB		] = 1.97
B10	0.26 (0.09)	b	b	b	b	b	b	b	b	b			
B10.A(4R)	0.27 (0.14)	k	k	b	b	b	b	b	b	b			] = 7.15
B10+B10.A(4R)	0.10 (0.08)										K, IA		
Microwell cultures 2 explants													
B10.BR	0.31 (0.15)												] = 5.82
B10.AKM	0.27 (0.14)												
B10.BR+B10.AKM	0.13 (0.07)										D only		

Data are from epithelial cultures grown as outgrowths from explants for 3 days in Hams F10 (modified with 0.02 M HEPES + 0.004 M NaHCO₃ pH 7.4) and 10% fetal calf serum. Contact inhibition was measured as nuclear overlap index on areas of $7.08 \times 10^{-8} \text{ m}^2$, 30 areas per culture type. $P = 0.01$ for $t = 2.66$, 58 d.f.

* Mean; s.d. in parentheses.

explanation of this effect would be that cells produce a soluble diffusible factor that affects the contact inhibition of cells allogeneic to the producer cells. Similar substances affecting cell adhesion have been described and characterised as glycoproteins from T and from B lymphocytes and have been shown to reduce the adhesion of allogeneic lymphocytes, if they are mismatched at H-2 D or L (T cells) or H-2 K (B cells)^{9,10}. There is some evidence for Ia antigen involvement in adhesion^{9,10}. It is likely that contact inhibition of movement is at least partially controlled by adhesion³.

It is not yet clear how these soluble factors are received by the target cells. Interestingly, Snell¹², reviewing work on MHC receptors in T cells, points out that the I region may contain pairs of closely linked genes determining interacting cell surface structures. If any locus is involved in adhesion, the presence of allogeneic cells (or their factors) may lead to competitive inhibition of such sites and thus diminution of adhesion.

Bodmer¹³ suggested that the true role of the MHC systems might be the control of cell-cell recognition. Normally, this recognition is, of course, predominantly a syngeneic phenomenon. The results we describe suggest that contact inhibition of movement, long believed to be involved in organogenesis, is under MHC control. Extreme contact inhibition seen in certain allogeneic combinations would prevent the interdigitation of cells required for the normal formation of tissues such as epithelia, and might, by itself, prevent proper fusion of allogeneic grafts. The problem yet to be solved is how syngeneic tissues within one organism differ sufficiently to separate and position themselves differently: one suggestion is that the display of MHC antigens varies between different tissues in such a way that pairs of cell types, one lacking and the other showing the expression of a certain locus, will interact in a different manner from those sharing its expression.

Received 25 May; accepted 10 August 1979.

1. Medawar, P. B. *Nature* **272**, 772 (1978).
2. Abercrombie, M. & Heaysman, J. E. M. *Expl Cell Res.* **6**, 293-306 (1954).
3. Abercrombie, M. *In Vitro* **6**, 128 (1970).
4. Abercrombie, M. *Archs Biol. (Liege)* **75**, 351 (1964).
5. Katz, D. *Lymphocyte Differentiation, Recognition and Regulation* (Academic, New York, 1977).
6. Zinkernagel, R. M. & Doherty, P. C. *Nature* **251**, 547-548 (1974).
7. Bartlett, P. F. & Edidin, M. *J. Cell Biol.* **77**, 377-388 (1978).
8. Zelensky, V., Matousek, V. & Lengerova, A. *J. Immunogenet.* **5**, 41-47 (1978).
9. Curtis, A. S. G. *J. Immunogenet.* **6**, 155-166 (1979).
10. Curtis, A. S. G. in *Cell Adhesion and Motility* (Cambridge University Press, in the press).
11. Middleton, C. A. *Expl Cell Res.* **70**, 91 (1972).
12. Snell, G. D. *Immun. Rev.* **38**, 3 (1978).
13. Bodmer, W. F. *Nature* **237**, 139-145; **237**, 183 (1972).

Mechanical activation and voltage-dependent charge movement in stretched muscle fibres

Chiu Shuen Hui* & W. F. Gilly†

Department of Physiology, Yale University Medical School, New Haven, Connecticut 06510

Although properties of the intramembranous charge movement in frog skeletal muscle have now been fairly well characterised¹⁻⁵, the function of this process remains uncertain⁶⁻⁹. Schneider and Chandler¹ first suggested that the observed charge movement in sartorius fibres held near slack length in a hypertonic paralysing solution took place within the transverse (T) tubular membranes in the region of the triad feet and signalled tubular depolarisation to the sarcoplasmic reticulum (SR), thereby initiating calcium release. More recently it has been reported that much of this charge movement is abolished in semitendinosus fibres stretched well beyond slack length in hypertonic Ringer solution¹⁰. Other studies using intracellular calcium indicators, however, have shown that substantial amounts of calcium are released by brief depolarisations in muscle fibres stretched to sarcomere lengths of 4.0 μm or more, when bathed in isotonic solutions¹¹⁻¹³. The question thus arises, is charge movement abolished in muscle fibres stretched in an isotonic Ringer solution? As calcium release persists in these conditions, an inhibition of charge movement would make involvement of this phenomenon with calcium release very unlikely. We report here the influence of stretch on charge movement and on an aspect of mechanical activation indicative of calcium release using standard, microelectrode voltage-clamp experiments on intact frog twitch fibres. We found no dramatic dependence of these properties on fibre length, and the possibility of a functional role for charge movement in the regulation of calcium release remains.

* Present address: Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907.

† To whom reprint requests should be addressed at: Department of Biology G/5, University of Pennsylvania, Philadelphia, Pennsylvania 19104.

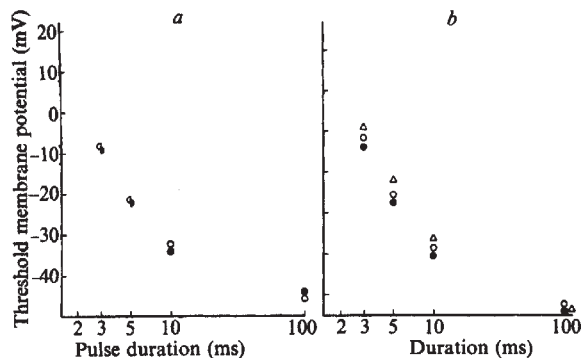


Fig. 1 Strength-duration curves for threshold mechanical responses in voltage-clamped twitch fibres at different sarcomere lengths. Membrane potential at threshold is plotted against duration of the depolarising voltage clamp pulse. Holding potential, -80 mV; temperature, 9 – 10 °C throughout. *a*, *Rana temporaria* sartorius muscle. Sarcomere lengths: $\circ$, 2.5 – 2.78 μm ; $\bullet$, 2.0 – 2.42 μm , mean values for eight fibres. *b*, *Rana pipiens*. $\bullet$, mean values from six sartorius fibres 2.27 – 2.50 μm long $\circ$, Δ : Means from six semitendinosus fibres at 2.97 – 3.07 μm ($\circ$) and 3.33 – 3.48 μm (Δ) (data obtained by L. L. Constantin on isolated, single fibres).

Threshold mechanical responses to relatively brief (<100 ms) depolarisations in a voltage-clamped muscle fibre are a functional, although indirect, indication of the voltage dependence of calcium release^{6,14–16}. The method has been described in detail elsewhere^{5,15}. A muscle fibre was visualised under a water immersion objective and impaled with two microelectrodes to measure the intracellular voltage and to pass current through a voltage clamp feedback loop. Sarcomere shortening in response to depolarising voltage clamp pulses of various durations was observed at the surface of the muscle fibre, and the membrane potential at which shortening was just detectable is defined as threshold. Mechanical threshold increases with decreasing pulse duration, so that a brief pulse must be much larger than a long one to bring the myoplasmic calcium level up to threshold.

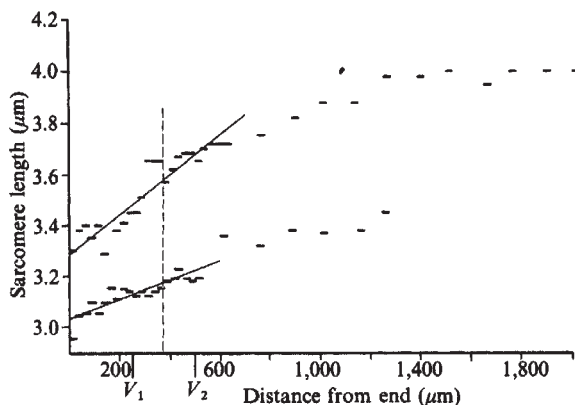


Fig. 2 Sarcomere length at the ends of stretched muscle fibres. Two fibres were studied in a semitendinosus muscle of *Rana pipiens* mounted exactly in the same way as for an electrophysiological experiment. The distance spanned by 10 sarcomeres was measured with a $\times 10$ Filar micrometer ($\times 40$ water immersion objective) and divided by 10 to yield sarcomere length. This value is plotted as a bar over the length of measurement. V_1 and V_2 indicate positions of the intracellular voltage sensing microelectrodes for charge movement experiments; current was sampled over the terminal 375 μm (vertical dotted line). Solid lines drawn by eye.

Figure 1a shows strength-duration (S – D) curves for threshold contractions in sartorius fibres from *Rana temporaria* at 9 °C. Means from two populations of fibres at different sarcomere lengths (2.0 – 2.4 μm and 2.5 – 2.8 μm) are shown. Long pulse thresholds represent a steady state balance between release and re-uptake of calcium by the SR, while short pulse

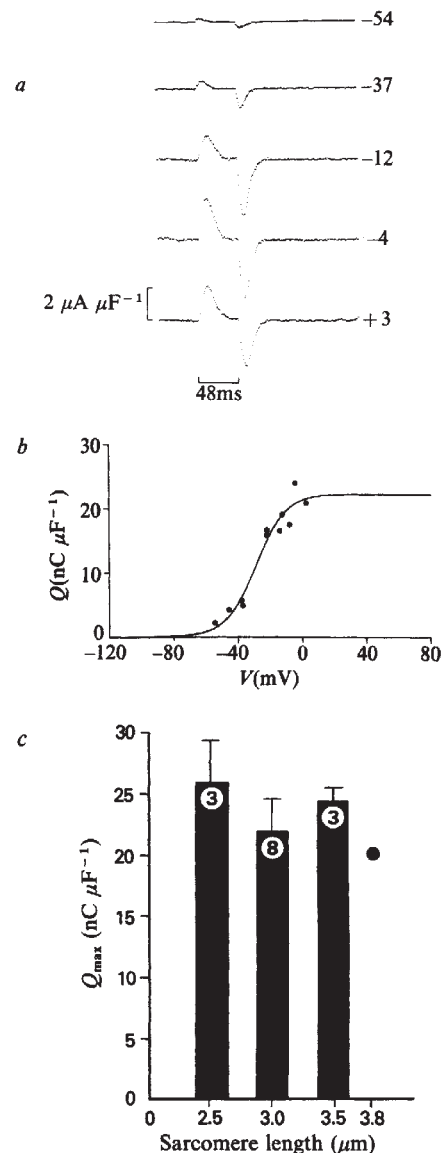


Fig. 3 Charge movement in stretched twitch fibres. *a*, Membrane current ($\Delta V = V_1 - V_2$) for depolarisation to the voltages, in mV, as indicated. Bathing solution contained 115 mM TEA-Cl, 2.5 mM KCl, 11.8 mM CaCl_2 , 1 mM PIPES, and 2 mM tetracaine. Electrodes spaced 250 μm apart (see Fig. 2); R_i assumed to be 255.2 $\Omega \text{ cm}$; 1 $\mu\text{A } \mu\text{F}^{-1}$ corresponds to 1.3 mV of ΔV . Fibre OCT23-78-F4; sarcomere length, 3.0 μm ; 7 °C; holding potential -80 mV; feedback control taken from V_2 . Uppermost trace is the average of eight sweeps; others are averages of four sweeps. See refs 2, 5 for details of measurement. *b*, Steady state Q – V plot for the fibre in (*a*). The points were obtained from the time integrals of current traces, some of which are shown in (*a*). The smooth curve was drawn according to the equation¹

$$Q(V) = Q_{\max} \left[1 + \exp \left(-\frac{V - \bar{V}}{k} \right) \right]^{-1}$$

and was fitted to the points by least squares. The best fit parameters were $\bar{V} = -28.6$ mV, $k = 9.3$ mV, $Q_{\max} = 22.1$ nC μF^{-1} . *c*, Maximum moveable, nonlinear charge, $Q_{\max}$, as a function of sarcomere length at the V_1 electrode. error bars indicate one standard deviation; number of fibres is indicated. Point at 3.8 μm is from one measurement at $+5$ mV.

thresholds (<10 ms) are probably set by the voltage-dependent kinetics of some regulatory step preceding calcium release. Values at all durations are indistinguishable in these two groups of fibres, suggesting no functional impairment of calcium release by this amount of stretch.

A somewhat greater length range was investigated in *Rana pipiens* sartorius and semitendinosus (single) fibres (Fig. 1b). Data from single fibres were obtained by the late Dr L. L. Costantin and have been tabulated previously¹⁵. Only the means for 2 ms of the 2.27–2.50 and 3.33–3.48 μm populations are significantly different at the 95% level by *t*-test. Again, a major influence of stretch on calcium release is not apparent.

To study membrane charge movements in intact fibres it is necessary to abolish contraction in response to depolarisation. We used an isotonic paralysing solution containing 2 mM tetracaine. Almers and Best⁶ have shown that this drug does not interfere with charge movement, and probably directly inhibits the calcium release process in the SR membranes. The solution also contained 115 mM tetraethylammonium (TEA) chloride and 5 mM RbCl to block K currents, tetrodotoxin to block Na currents, 11.8 mM CaCl_2 to minimise leakage conductance around the electrodes, and 1 mM PIPES buffer (pH 7.0). Temperature was 7 °C.

The three electrode voltage-clamp technique used allows an approximate measurement of membrane current density at the end of a muscle fibre¹⁷. Feedback control was taken from the electrode more distant (500 μm) from the fibre end (V_2), and the current trace was integrated to determine capacitance as described by Schneider and Chandler¹⁸. Pulse duration was 48 ms; control pulses were obtained over the range of –80 to –150 mV (refs 1, 2). Membrane current amplitude and the quantity of charge were normalised by the effective membrane capacitance, to be in units of $\mu\text{A } \mu\text{F}^{-1}$ and $\text{nC } \mu\text{F}^{-1}$, respectively. This removes the dependence of charge movement measurements on fibre radius, internal resistivity (R_i), and inter-electrode spacing.

Inhomogeneity of sarcomere spacing at the fibre's end¹⁹ adds a complication to these experiments, however. As current is effectively sampled over the last several hundred micrometres of fibre length, a large number of short sarcomeres at the very end might mask any effect of stretch. To check this, sarcomere spacing was measured at the ends of two semitendinosus fibres in a bundle pinned out in the same manner as for a voltage-clamp experiment. Results are shown in Fig. 2. In both cases a substantial difference in sarcomere length does exist between the end and the middle of the fibre. Over the region in which current was measured in these experiments (the terminal 375 μm), however, there is only a 5–10% decrease in sarcomere length. Charge movement measurements are referred to the sarcomere length measured at the V_1 electrode, 250 μm from the fibre end. It is not likely that an inhomogeneity of sarcomere spacing in this region seriously affects our results.

Figure 3a shows examples of charge movement traces at several voltages from a semitendinosus twitch fibre stretched to 3.0 μm sarcomere length. The magnitude of the charge movement increases and the kinetics become faster as the pulses get larger. Figure 3b shows the amount of charge moved during each pulse. The solid curve is drawn according to the usual Boltzmann distribution model¹ with a steepness of *e*-fold for 9 mV and a saturating Q_{max} value of 22 $\text{nC } \mu\text{F}^{-1}$. The shape of this *Q*–*V* curve is very similar to that obtained from fibres at rather shorter lengths in both isotonic⁶ and hypertonic^{2,3} solutions, although hypertonicity appears to shift the curve 15–20 mV in the negative direction⁶.

Figure 3c summarises results from semitendinosus twitch fibres at different degrees of stretch. The total amount of moveable charge, Q_{max} , does not significantly depend on sarcomere length. The point at 3.8 μm is for a single pulse to +5 mV, immediately after which the fibre broke.

While Q_{max} and the steepness of the *Q*–*V* distribution do not appear to depend greatly on stretch (see also ref. 20), these are steady state properties, and it remains possible that stretch does

slow the kinetics of charge movement. This could be due to a real effect of stretch on the process generating the charge movement or to an indirect effect on some cable properties of the muscle fibre. The records in Fig. 3a are not obviously slower than other published records; additional experiments are required to study more carefully the effect of stretch on charge movement kinetics. Any effect of stretch on kinetics might also be augmented in a hypertonic Ringer solution.

Neither the properties of charge movement nor of mechanical activation we measured in fibres in isotonic solutions are significantly affected by stretch to sarcomere lengths at which the contractile filaments no longer overlap. These observations are consistent with the possibility that charge movement in skeletal muscle has a role in regulating calcium release.

We thank Dr S. M. Baylor for discussions and the 3.8- μm fibre. This work was supported by a grant from the NIH to Dr W. K. Chandler.

Received 31 May; accepted 2 August 1979.

1. Schneider, M. F. & Chandler, W. K. *Nature* **242**, 244–246 (1973).
2. Chandler, W. K., Rakowski, R. F. & Schneider, M. F. *J. Physiol., Lond.* **254**, 245–283 (1976).
3. Adrian, R. H. & Almers, W. *J. Physiol., Lond.* **254**, 339–360 (1976).
4. Almers, W. *J. Physiol., Lond.* **262**, 613–637 (1976).
5. Gilly, W. F. & Hui, C. S. *J. Physiol., Lond.* (in the press).
6. Almers, W. & Best, P. M. *J. Physiol., Lond.* **262**, 583–611 (1976).
7. Chandler, W. K., Rakowski, R. F. & Schneider, M. F. *J. Physiol., Lond.* **254**, 285–316 (1976).
8. Adrian, R. H. & Peres, A. *J. Physiol., Lond.* **289**, 83–97 (1979).
9. Chandler, W. K., Gilly, W. F. & Hui, C. S. in *Biophysical Aspects of Cardiac Muscle* (ed. Morad, M.) 31–44 (Academic, New York, 1978).
10. Adrian, R. H., Caputo, C. & Huang, C. L.-H. *J. Physiol., Lond.* **284**, 151P (1978).
11. Miledi, R., Parker, I. & Schalow, G. *Proc. R. Soc. B* **198**, 201–210 (1977).
12. Baylor, S. M., Chandler, W. K. & Marshall, M. W. *Biophys. J.* **25**, 141a (1979).
13. Rios, E. L., Kovacs, L. & Schneider, M. F. *Biophys. J.* **25**, 201a (1979).
14. Adrian, R. H., Chandler, W. K. & Hodgkin, A. L. *J. Physiol., Lond.* **204**, 207–230 (1969).
15. Costantin, L. L. *J. gen. Physiol.* **63**, 657–674 (1974).
16. Gilly, W. F. & Hui, C. S. *Nature* **266**, 186–188 (1977).
17. Adrian, R. H., Chandler, W. K. & Hodgkin, A. L. *J. Physiol., Lond.* **208**, 607–644 (1970).
18. Schneider, M. F. & Chandler, W. K. *J. gen. Physiol.* **67**, 125–163 (1976).
19. Huxley, A. F. & Peachey, L. D. *J. Physiol., Lond.* **156**, 150–163 (1961).
20. Hui, C. S. & Gilly, W. F. *Biophys. J.* **25**, 200a (1979).

Renal brush border membrane adaptation to phosphorus deprivation in the *Hyp/Y* mouse

H. S. Tenenhouse & C. R. Scriver

Medical Research Council Genetics Group, Human Genetics Center, (McGill University) and De Belle Laboratory for Biochemical Genetics, McGill University-Montreal Children's Hospital, Research Institute, 2300 Tupper Street, Montreal, Quebec, Canada H3H 1P3

Human X-linked hypophosphataemia (XLH)¹ involves a partial loss in net renal reabsorption of phosphate anion (P_i). The mutant male hemizygote has about half the normal renal transport activity² and this is quite insensitive to circulating parathyroid hormone (PTH)^{2,3}. The renal defect is sufficient to explain persistent hypophosphataemia in XLH. The hypophosphataemic mouse (*Hyp*)⁴ also has an X-linked defect in renal reabsorption of P_i (refs 4–7). The mutant male hemizygote (*Hyp/Y*) has about half the normal renal reabsorption activity and this abnormality persists after parathyroidectomy⁷. Conservation of the X chromosome during mammalian evolution and evidence for functional homology between species for specific X-linked genes⁸ have encouraged us to use the *Hyp* mouse as a model for the renal defect in XLH. We have studied P_i transport in purified renal brush border membrane vesicles (BBMV) prepared from *Hyp/Y* mice^{5,9}. A selective loss involving almost half the Na^+ -dependent P_i co-transport activity is

present in *Hyp/Y* membranes. These findings, genetic studies in man^{2,10} and recent physiological evidence¹¹ suggest that P_i transport in the mammalian nephron occurs by at least two processes: one under the control of an X-linked gene that is more sensitive to PTH inhibition and another process, controlled by an autosomal locus, that is less sensitive to the hormone. We report here on the ability of the *Hyp/Y* mouse to adapt to phosphorus deprivation; the response is observed both *in vivo* in net reabsorption of P_i and *in vitro* in BBMV transport of P_i .

Hyp mice were bred and reared in our laboratory. Adult (>100 d old) *Hyp/Y* and normal *+Y* littermates were given either a control diet containing 0.4% (w/w) phosphorus or a low-P diet containing 0.03% phosphorus. Both diets contained the same amount of calcium (0.4% w/w). The animals were not thyroparathyroidectomised. After 2 weeks, mice were placed in Gelman metabolic cages; blood and urine was collected and analysed for calcium, phosphorus and creatinine content measured as before⁵.

Table 1 shows that in both *Hyp/Y* and *+Y* mice there was *in vivo* renal adaptation to phosphorus deprivation, whereas calcium excretion increased, as expected¹². Renal adaptation did not fully correct the mutant phenotype: *Hyp/Y* animals, in the adapted state, maintained lower ($P < 0.001$) plasma levels and a higher ($P < 0.001$) fractional excretion of phosphorus than *+Y* mice in the adapted state.

Purified BBMVs were prepared from *Hyp/Y* and *+Y* renal cortex by the method of Booth and Kenny¹³ and incubated as

before^{5,9}. Net uptake of $^{32}P_i$ (100 μ M) by BBMVs was measured after 15 or 30 s incubation in the presence of NaCl or KCl gradients. The simultaneous uptake of D-glucose (20 μ M) was also measured and Na^+ /glucose co-transport was used as an internal control. D-glucose and P_i at their respective concentrations did not interfere with the co-transport of each other, and the possibility of significant binding of phosphate to BBMVs was excluded by the appropriate osmotic experiments in a sucrose medium⁹.

P_i transport activity, both absolute and normalised to simultaneous D-glucose uptake, was significantly greater ($P < 0.001$) in BBMVs prepared from *Hyp/Y* and *+Y* mice adapted to phosphorus deprivation than in BBMVs prepared from animals fed the control diet (Table 2). The characteristic deficiency of Na^+ / P_i co-transport activity in *Hyp/Y* membranes was retained in the adapted state. We also examined the kinetics of P_i uptake by adapted BBMVs, and found that although there was no change in the apparent K_m value, the V_{max} for uptake was increased twofold in the adapted state (Fig. 1).

These findings are of interest for the interpretation of *in vivo* adaptation to phosphorus deprivation, a phenomenon reported earlier for rat¹⁴ and dog¹⁵. Others have observed that the adapted kidney becomes insensitive to PTH¹⁶, that adaptation is not dependent on the integrity of parathyroid or thyroid glands^{17,18} and that, depending on the species, days or weeks are required for the phenomenon to appear^{15,18}. The adaptation is a function of the early proximal renal tubule although more distal parts may also be involved^{12,19}. Our findings indicate that renal

Table 1 *In vivo* renal adaptation to chronic phosphorus deprivation in *+Y* and *Hyp/Y* littermates (mean $\pm$ s.e.)

	<i>+Y</i> mice		<i>Hyp/Y</i> mice	
	Control diet (n = 4)	Low-P diet (n = 4)	Control diet (n = 4)	Low-P diet (n = 4)
Plasma P_i (mg dl ⁻¹)	8.33 $\pm$ 0.62	5.09 $\pm$ 0.84*	3.97 $\pm$ 0.19†	1.01 $\pm$ 0.14*†
Phosphorus excretion index	0.280 $\pm$ 0.021	0.005 $\pm$ 0.001*	0.569 $\pm$ 0.024†	0.034 $\pm$ 0.004*†
Calcium excretion index	0.015 $\pm$ 0.0002	0.175 $\pm$ 0.003*	0.016 $\pm$ 0.0003	0.192 $\pm$ 0.004*

Littermate animals were classified as *+Y* (normal male) and *Hyp/Y* (mutant hemizygote) according to quantitative and qualitative phenotypic features⁴. Adult mice were given specified diets (ICN) for 2 weeks. Urine was collected for about 3 h, after which blood was obtained from the tail. Samples were analysed as described elsewhere⁵. Total phosphorus content of diets: control, 0.4% (w/w); Low-P, 0.03% determined by chemical analysis according to Trohler *et al.*¹⁷. Excretion index (analogous to fractional excretion) is derived from the equation: (urine [P_i] or [Ca] $\times$ plasma [creatinine]) / (plasma [P_i] or [Ca] $\times$ urine [creatinine]).

*Different from control-diet animals: $P < 0.001$, Student *t*-test.

†Different from *+Y* animals on corresponding diet: $P < 0.001$, Student *t*-test.

Table 2 Uptake of ^{32}P and ^{14}C -D-glucose by renal brush border membrane vesicles from adult control (*+Y*) and hypophosphataemic (*Hyp/Y*) male mice

	<i>+Y</i>		<i>Hyp/Y</i>	
	Control diet (0.4% P)	Low-P diet (0.03% P)	Control diet (0.4% P)	Low-P diet (0.03% P)
P_i uptake (30 s) (pmol per mg protein)				
In KCl	81.4 $\pm$ 7.4	67.8 $\pm$ 8.9	60.7 $\pm$ 8.6	55.7 $\pm$ 3.8
In NaCl	807.8 $\pm$ 55.8	1,596.2 $\pm$ 84.7*	617.4 $\pm$ 31.6†	902.5 $\pm$ 82.3†*
Normalised P_i uptake (30 s; Na dependent) (pmol P per pmol D-glucose)	9.0 $\pm$ 0.2	25.2 $\pm$ 1.4*	3.6 $\pm$ 0.2†	14.5 $\pm$ 0.7†*

Net uptake of ^{32}P -labelled orthophosphate (100 μ M) was measured at 30 s by BBMVs prepared from renal cortex of normal male (*+Y*) and hemizygous (*Hyp/Y*) littermates. The membranes were isolated and purified by the method Booth and Kenny¹³, and suspended in 300 mM mannitol and 20 mM Tris-HEPES, pH 7.4, at a final protein concentration of 2–3 mg ml⁻¹. Uptake was measured by the Millipore technique used previously^{5,9}. The reaction was initiated by adding 90 μ l of incubation medium containing ^{32}P and ^{14}C -D-glucose (20 μ M) to 10 μ l of brush border membrane preparation at 20°C. The reaction was terminated by the addition of 4 ml stop solution (100 mM mannitol, 100 mM NaCl, 20 mM Tris-HEPES and 10 mM arsenate pH 7.4, 4°C), and filtered immediately through a 0.45- μ m Millipore filter. After two rapid additional washes with stop solution the filters were counted in Aquasol-2 in a liquid scintillation counter. Incubation solutions were: (1) 100 mM KCl, 100 mM mannitol and 20 mM Tris-HEPES, pH 7.4; (2) 100 mM NaCl, 100 mM mannitol and 20 mM Tris-HEPES, pH 7.4. Background counts were subtracted from experimental counts; the former were always less than 15% of uptake at equilibrium in the presence of membranes. Phosphate is not significantly bound to BBMVs and is entirely retained as the inorganic anion after uptake^{5,9}.

*Difference between low P and control diet is significant: $P < 0.001$, Student *t*-test.

†Difference between *+Y* and *Hyp/Y* is significant: $P < 0.001$, Student *t*-test. Data are mean $\pm$ s.e.m. of five determinations.

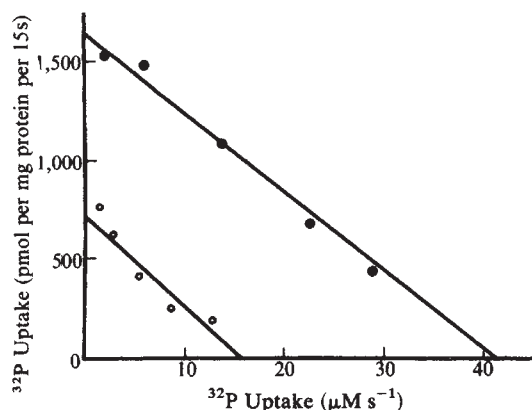


Fig. 1 Eadie-Hofstee plots of the Na^+ -dependent uptake of ^{32}P at 15 s by BBMVs prepared from kidneys of adult normal mice fed control diet (0.4%, w/w) (○) and low-P diet (0.03%) (●), each for 2 weeks. The apparent K_m values are $45 \mu\text{M}$ in mice fed the control diet and $40 \mu\text{M}$ in P-deprived mice; V_{max} is increased in the adapted animal (1,659 pmol per mg protein per 15 s compared with 722 pmol per mg protein per 15 s ($P < 0.001$)). The correlation coefficients for the regressions were 0.886 and 0.850 for control and P-deprived animals, respectively, over the concentration range 15–800 μM P_i . See Table 2 for details of BBMVs incubation and transport assay. Kinetic values refer to Na^+ -dependent uptake exclusively, that is, total uptake in NaCl minus diffusional uptake in KCl .

adaptation to phosphorus deprivation *in vivo* involves a change in P_i transport activity in the brush border membrane. This adaptation occurs in the presence of parathyroid hormone *in vivo* and seems to involve a selective change in the number or activity of carriers specific for phosphate anion. Because adaptation occurs in the *Hyp/Y* animal, we suggest that a brush border membrane mechanism for P_i transport, other than that controlled by the X-linked gene, is responsible for the enhanced BBMVs P_i transport activity. The signal for this adaptation is unknown: the hypophosphataemic *Hyp/Y* phenotype itself is insufficient to produce this, and the additional stimulus of environmental phosphorus deprivation is required to evince the membrane responsible in the *Hyp* animal. It is interesting that renal phosphorus content in *Hyp* mice does not differ from normal⁵. We have also observed that phosphorus levels in renal cortex are similar in control and adapted mice²⁰.

This work was supported by the MRC of Canada (Genetics Group Award). *Hyp* breeding pairs were obtained from Dr Eva Eicher.

Received 26 April; accepted 1 August 1979.

- Rasmussen, H. & Anast, C. in *The Metabolic Basis of Inherited Disease* 4th edn (eds Stanbury, J. B., Wyngaarden, J. B. & Fredrickson, D. S.) 1537–1562 (McGraw-Hill, New York, 1978).
- Glorieux, F. & Scriver, C. R. *Science* **175**, 997–1000 (1972).
- Scriver, C. R., Glorieux, F. H., Reade, T. M. & Tenenhouse, H. S. in *Inborn Errors of Calcium and Bone Metabolism* (eds Bickel, H. & Stern, J.) 150–178 (University Park Press, Baltimore, 1976).
- Eicher, E. M., Southard, J. L., Scriver, C. R. & Glorieux, F. H. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4667–4671 (1976).
- Tenenhouse, H. S., Scriver, C. R., McInnes, R. R. & Glorieux, F. H. *Kidney Int.* **14**, 236–244 (1978).
- Giasson, S. D., Brunette, M. G., Danan, G., Vigneault, N. & Carriere, S. *Pflügers Arch. ges. Physiol.* **371**, 33–38 (1977).
- Cowgill, L., Goldfarb, S., Lau, K., Slatopolsky, E. & Agus, Z. S. *J. clin. Invest.* **63**, 1203–1210 (1979).
- Ohno, S. *The Sex Chromosomes and Sex-Linked Genes*, 46–73 (Springer, New York, 1967).
- Tenenhouse, H. S. & Scriver, C. R. *Can. J. Biochem.* **56**, 640–646 (1978).
- Scriver, C. R., MacDonald, W., Reade, T., Glorieux, F. H. & Nogrady, B. *Am. J. med. Genet.* **1**, 101–117 (1977).
- Dennis, V. M., Bello-Reusa, E. & Robinson, R. R. *Am. J. Physiol.* **233**, F29–F38 (1977).
- Goldfarb, S., Westby, G. R., Goldberg, M. & Agus, Z. S. *J. clin. Invest.* **59**, 770–779 (1977).
- Booth, A. G. & Kenny, A. J. *Biochem. J.* **142**, 575–581 (1974).
- Steele, T. H. & DeLuca, H. F. *J. clin. Invest.* **57**, 867–874 (1976).
- Wen, S.-F., Boynar, J. W. Jr & Stoll, R. W. *Am. J. Physiol.* **3**, F199–F206 (1978).
- Steele, T. H., Underwood, J. L., Stromberg, B. A. & Larmore, C. A. *J. clin. Invest.* **58**, 1461–1464 (1976).
- Tröhler, U., Bonjour, J.-P. & Fleisch, H. *J. clin. Invest.* **57**, 264–273 (1976).
- Tröhler, U., Bonjour, J.-P. & Fleisch, H. *Nature* **261**, 145–146 (1976).
- Mühlbauer, R. C., Bonjour, J.-P. & Fleisch, H. *Am. J. Physiol.* **233**, F342–F348 (1977).
- Tenenhouse, H. S. & Scriver, C. R. *Can. J. Biochem.* **57**, 938–944 (1979).

Is adenosine the mediator of opiate action on neuronal firing rate?

Trevor W. Stone & Martin N. Perkins

Department of Physiology, St. George's Hospital Medical School, University of London, London SW1, UK

Despite recent advances^{1,2}, relatively little is known of the mechanism of action of opiate analgesics at the cellular level. There is evidence, however, that the inhibition of neurotransmitter release produced by morphine may be mediated by the initial release of adenosine^{3–5}; the effect is mimicked by adenosine and the actions of both adenosine and morphine are blocked by theophylline and enhanced by dipyridamole^{3–5}. When applied microiontophoretically, adenosine has also been shown to be a potent depressant of the firing rate of single neurones in the mammalian brain, and a similar depressant effect of morphine has been claimed to be stereospecific, antagonised by naloxone^{6–9} and greatly reduced in animals made tolerant to, and dependent on, morphine^{8,9}. The experiments described here examine the effect of aminophylline, a soluble theophylline derivative, on these depressant effects. We have found that aminophylline will block responses to both morphine and adenosine, suggesting that the depressant responses to morphine may also be mediated by the local release of adenosine.

Male Porton Wistar rats were anaesthetised with urethane, 1.3 g per kg, intraperitoneally. The animals were placed in a stereotaxic frame and the body temperature automatically maintained at 37–38 °C. A burr hole was made so as to allow electrode penetration into the corpus striatum at coordinates according to the atlas of DeGroot¹⁰.

Drugs were applied by microiontophoresis from six- or seven-barrelled micropipettes of tip diameters 4–12 μm . The drugs used were: morphine sulphate, 50 or 100 mM, pH 5.5; adenosine hemisulphate, 100 mM, pH 4; γ -aminobutyric acid (GABA), 100 mM, pH 4; naloxone HCl 50 mM, pH 5; aminophylline (theophylline ethylenediamine), 50 mM, pH 9 (unadjusted). The latter drug was injected as an anion, and the rest as cations. No retaining current was applied to aminophylline. One barrel of the micropipette contained 165 mM NaCl for current balancing and testing. Action potentials were recorded through another barrel of the multibarrel pipette, filled with 3 M NaCl, or through a separate, single microelectrode glued alongside the multibarrel¹¹. The spikes were amplified, gated and displayed on oscilloscopes and the gated pulses were counted and displayed on a polygraph as a record of spikes per s.

The effect of local application of morphine was studied in 127 neurones in the rat striatum using currents of 60–160 nA. Of these, 72 were found to be depressed, 23 excited and 32 did not respond to the alkaloid. The firing of most of the inhibited units was suppressed completely within 1–2 s and recovery from this depression was also rapid, the original firing rate returning within 5–10 s of terminating the ejection current (Fig. 1). On 20 quiescent cells morphine and adenosine also reduced the excitation induced by iontophoretically applied glutamate. The time course of the depression was similar to that of spontaneously firing units.

Adenosine and GABA also depressed neurones responding to morphine (Fig. 1) when applied with currents ranging from 45 to 120 nA. Again, the onset and recovery from these inhibitions were rapid, as has been described previously¹².

The opiate antagonist naloxone was then applied to 29 cells in an attempt to block the responses to morphine. Typically, both inhibitory and excitatory responses were difficult to antagonise, a problem exacerbated by the direct depressant effect of naloxone itself. However, 14 cells did show a reduction of morphine responses whereas GABA responses were unaffected, although recovery from naloxone blockade was slow.

The effects of aminophylline on morphine responses has been tested on 15 neurones. Nine of these exhibited some increase of

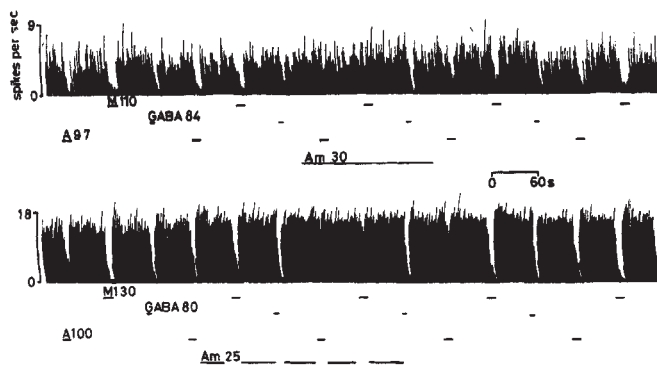


Fig. 1 Ratemeter records of the firing rate of two spontaneously active neurones in the rat striatum. The firing of both cells was depressed by the application by microiontophoresis of adenosine (A), morphine (M) and GABA. The application of aminophylline (Am) as an anion rapidly blocked the responses to adenosine and morphine, whereas GABA responses were unaffected. The times of drug application are shown by bars below the records, and the figures are the iontophoretic currents used, in nA. Time bar, 60 s.

baseline firing rate during the ejection of aminophylline. In all 15 cells the depressant responses to morphine were rapidly and reversibly antagonised by doses of aminophylline ranging from 14 to 90 nA, and an example of this is shown in Fig. 1. Depressant responses to adenosine were always greatly reduced or abolished in parallel with the loss of morphine responses, whereas the effects of GABA remained unaffected (Fig. 1). Ethylenediamine had no effect on the agonist responses.

These experiments demonstrate a powerful depression by morphine of the firing rate of neurones in the rat corpus striatum, an area which was chosen for this study because it contains a high density of opiate receptors^{13,14}. The ability of naloxone to antagonise the morphine depression of firing supports the view expressed by others that such depressant responses result from an action on the specific opiate receptor⁶⁻⁹, although the difficulty experienced with naloxone may result from an action of morphine on several classes of receptor, as discussed by Lord *et al.*¹⁵.

The inhibitory effect of iontophoretically applied adenosine has been previously described as a widespread phenomenon in the central nervous system^{12,16-19}. Furthermore, the effects of this purine in many tissues and preparations can be blocked by methylxanthine derivatives such as theophylline^{3-6,12,16-22}. Our demonstration that aminophylline can antagonise responses to both adenosine and morphine but not GABA, while the effects of morphine, but not adenosine, can be blocked by naloxone leads us to propose that the first result of the interaction of morphine with its receptor is a release of adenosine from the tissue. Such a sequence presumably also explained the inhibition of transmitter release by morphine and its modulation by dipyridamole and theophylline reported by other workers using peripheral^{3,4} or central⁵ models of transmitter release. Indeed, Fredholm and Vernet²³ have now shown directly that morphine will enhance the depolarisation-induced release of adenosine from neurones²³.

It could be argued that, as adenosine and morphine are potent inhibitors of calcium-dependent neurotransmitter release, such an action could contribute to their depressant effects on neuronal firing. Aminophylline might then reduce that action by increasing calcium availability in the presynaptic terminals. However, the only relevant work seems to be that of Edstrom and Phillis²⁴, who reported that adenosine hyperpolarised cortical neurones without affecting membrane conductance, although this finding, as the authors noted, does not prove a presynaptic site of action. On the other hand, we have demonstrated that the depressant effects of morphine and adenosine occur against excitatory responses to glutamate, an amino acid thought to act postsynaptically by increasing sodium conductance²⁵.

We also suggest that the release of adenosine may be highly relevant to the analgesic effects of morphine in the conscious animal, as morphine analgesia has been shown to be reduced by the systemic administration of theophylline²⁶. Furthermore, Horrobin has claimed that adenosine is a potent antagonist of many of the effects of prostaglandins^{27,28}. A release of adenosine by opiates as suggested here may therefore explain the antagonism by morphine of prostaglandin-induced increases in cyclic AMP^{29,30}, an action which Collier has proposed to be central to an understanding of opiate tolerance and dependence³¹.

This work was supported by the MRC and Action Research for the Crippled Child.

Received 4 May; accepted 13 August 1979.

1. Kosterlitz, H. W. (ed.) *Opiates and Endogenous Opioid Peptides* (North-Holland, Amsterdam, 1976).
2. Hughes, J. (ed.) *Centrally Acting Peptides* (Macmillan, London, 1978).
3. Gintzler, A. R. & Musacchio, J. M. *J. Pharmac. exp. Ther.* **194**, 575-582 (1975).
4. Jhamandas, K., Sawynok, J. & Sutak, M. *Eur. J. Pharmac.* **49**, 309-312 (1978).
5. Sawynok, J. & Jhamandas, K. H. *J. Pharmac. exp. Ther.* **197**, 379-390 (1976).
6. Bradley, P. B., Briggs, I., Gayton, R. J. & Lambert, R. A. *Nature* **261**, 425-426 (1976).
7. Young, W. S., Bird, S. J. & Kuhar, M. J. *Brain Res.* **129**, 366-370 (1977).
8. Satoh, M., Zieglängsberger, W., Fries, W. & Herz, A. *Brain Res.* **82**, 378-382 (1974).
9. Zieglängsberger, W., Fry, J. P., Herz, A., Moroder, L. & Wunsch, E. *Brain Res.* **115**, 160-164 (1976).
10. DeGroot, T. *Trans. R. Neth. Acad. Sci.* **52**, 1-40 (1959).
11. Stone, T. W. *J. Physiol., Lond.* **233**, 211-225 (1973).
12. Stone, T. W. & Taylor, D. A. *J. Physiol., Lond.* **266**, 523-543 (1977).
13. Pert, C. B. & Snyder, S. H. *Science* **179**, 1011-1014 (1976).
14. Goldstein, A. *Life Sci.* **14**, 615 (1974).
15. Lord, J. A. H., Waterfield, A. A., Hughes, J. & Kosterlitz, H. W. *Nature* **267**, 495-499 (1977).
16. Phillis, J. W., Kostopoulos, G. K. & Limacher, J. J. *Can. J. Physiol. Pharmac.* **52**, 1226-1229 (1974).
17. Kostopoulos, G. K. & Phillis, J. W. *Exptl Neurol.* **55**, 719-724 (1977).
18. Taylor, D. A. & Stone, T. W. *Experientia* **34**, 481-482 (1978).
19. Stone, T. W. *Biochem. Soc. Trans.* **6**, 858-862 (1978).
20. Sattin, A. & Rall, T. W. *Molec. Pharmac.* **6**, 13-23 (1970).
21. Blume, A. J., Dalton, C. & Sheppard, H. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3099-3102 (1973).
22. Green, R. D. & Stanberry, L. R. *Biochem. Pharmac.* **26**, 37-43 (1977).
23. Fredholm, B. B. & Vernet, L. *Acta. physiol. scand.* **104**, 502-504 (1978).
24. Edstrom, J. P. & Phillis, J. W. *Can. J. Physiol. Pharmac.* **54**, 787-790 (1976).
25. Zieglängsberger, W. & Puil, E. A. *Exptl Brain Res.* **17**, 35-49 (1973).
26. Ho, I. K., Loh, H. H. & Leong Way, E. *J. Pharmac. exp. Ther.* **185**, 336-346 (1973).
27. Ally, A. I. *et al. Prostaglandins* **14**, 109-118 (1977).
28. Karmali, R. A. *et al. Res. Commun. chem. Path. Pharmac.* **19**, 181-184 (1978).
29. Collier, H. O. J. & Roy, A. C. *Nature* **248**, 24-27 (1974).
30. Havemann, U. & Kuschinsky, K. *Archs pharmac.* **302**, 103-106 (1978).
31. Collier, H. O. J., Francis, D. L. & Roy, A. C. *Adv. Biochem. Psychopharmac.* **15**, 337-345 (1976).

Vanadate inhibits uncoupled Ca efflux but not Na-Ca exchange in squid axons

R. DiPolo*, H. R. Rojas* & L. Beaugé†

* Laboratorio de Permeabilidad Iónica, Instituto Venezolano de Investigaciones Científicas, Caracas, Venezuela

† Physiological Laboratories, Cambridge, UK

Nerve cells can maintain a very low intracellular calcium concentration ($[Ca^{2+}]_i$) against large Ca^{2+} electrochemical gradients (see ref. 1 for review). The properties of the calcium efflux from these cells depend on $[Ca^{2+}]_i$ (ref. 2), and within the physiological range, most Ca efflux depends on ATP (which stimulates with high affinity) and is insensitive to Na_i , Na_o and Ca_o (uncoupled Ca efflux). When the $[Ca^{2+}]_i$ is well above the physiological range, Ca efflux becomes only partially dependent on ATP (acting now with low affinity), is inhibited by Na_i and is stimulated by Na_o and Ca_o (Na-Ca exchange). Orthovanadate, a powerful inhibitor of the $(Na^+ + K^+)ATPase$ and the Na pump^{3,4}, also inhibits the Ca-stimulated ATPase activity, which is the enzymatic basis for the uncoupled Ca pump, in human red cells⁵. The experiments reported here show that in squid axons the ATP-dependent uncoupled Ca efflux can be fully and reversibly inhibited by vanadate, whereas concentrations of vanadate 10 times higher have no effect on the Na-Ca exchange. This is another indication that the uncoupled Ca efflux represents an ATP-driven Ca pump, and supports the suggestion that the uncoupled Ca efflux and Na-Ca exchange are mediated by different mechanisms².

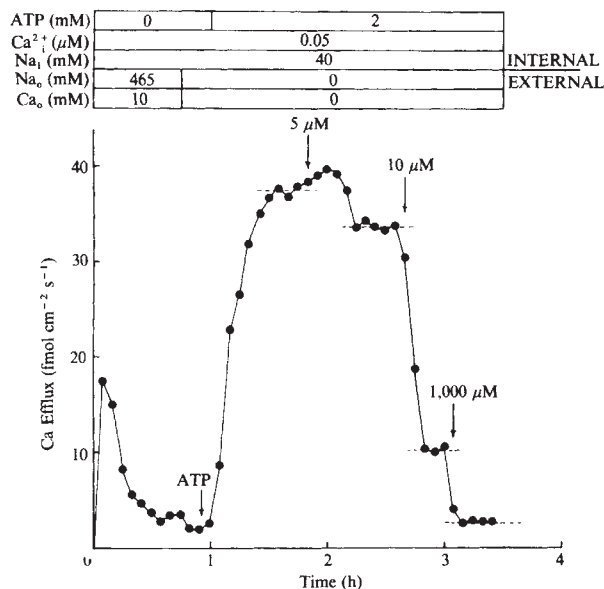


Fig. 1 Effect of different internal vanadate concentrations on Ca efflux in an axon dialysed with 0.05 μM [Ca²⁺]_i (axon 150379, 400 μM). The composition of the dialysis solution was (mM): K⁺, 310; Na⁺, 35; Mg²⁺, 4 in excess to the ATP concentration; Tris⁺, 5; Cl⁻, 79; aspartate, 310; EGTA, 1; glycine, 330; and oligomycin, 10 μg ml⁻¹; at 980 mosmol, pH 7.1 (17 °C). ATP (Boehringer) and phosphoarginine (PA) (Calbiochem) were used as sodium salts and were stored at -90 °C as stock 250 mM solutions at pH 7.1 adjusted with Tris-OH (ATP) and HCl (PA). The ATP solution also contained 250 mM MgCl₂. The composition of the standard artificial seawater was (mM): Na⁺, 465; K⁺, 10; Mg²⁺, 50; Ca²⁺, 10; Tris⁺, 10; Cl⁻, 580; EDTA, 0.1; 1,000 mosmol, pH 7.6 (17 °C). Tris was used as a substitute for Na. Removal of Na_o and Ca_o was carried out without changing the other constituents. The nominally Ca-free seawater also contained 0.5 mM EGTA and no added Ca. The axon was incubated for 1 h before dialysis in artificial seawater containing 1 mM KCN. The arrows refer to addition of vanadate and the concentration is indicated. The flow of the dialysis solution was 1 μl min⁻¹. The flow of external solutions was about 1 ml min⁻¹ and was collected at 5-min intervals, added to 5 ml of scintillator and counted. For more details of the dialysis technique see ref. 6.

The experiments were carried out on giant axons taken from live specimens of the tropical squid *Dorytheutis plei*. The intracellular concentration of low molecular weight solutes was controlled using the internal dialysis technique⁶. To inhibit endogenous ATP production, the axons were soaked in seawater containing 1 mM KCN for 1 h before the dialysis began. The production of ADP from the hydrolysis of exo-

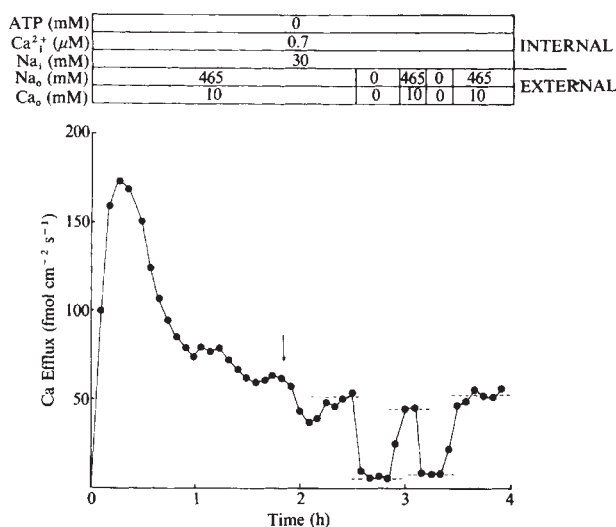


Fig. 2 The lack of effect of 1 mM vanadate applied internally (arrow) on the Na-Ca exchange in squid giant axon (axon 190179 B, 450 μM). In this experiment the axon was dialysed with a solution lacking ATP (to inhibit the uncoupled Ca efflux) and with 0.7 μM [Ca²⁺]_i to maximise the Na-Ca exchange component. The general technique was the same as that described in Fig. 1 legend.

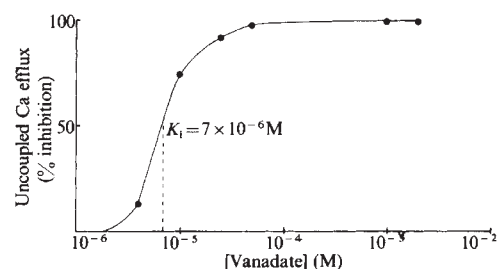


Fig. 3 Dose-response curve for the vanadate inhibition of the ATP-dependent 'uncoupled' Ca efflux in squid giant axons. The general technique was similar to that in Fig. 1, where after a steady Ca efflux was obtained in the presence of ATP, different concentrations of vanadate were tested. In all cases the external solutions lacked both Na and Ca. Each point is the mean of three axons. The inhibitory effect of vanadate was always reversible. ATP, 2 mM; Na_o, 40 mM; K_o, 10 mM; 17 °C.

ogenous ATP was minimised by adding 5 mM phosphoarginine to the dialysis solutions.

Figure 1 shows that in an axon dialysed with 0.05 μM Ca²⁺ the ATP-dependent 'uncoupled' Ca efflux is progressively inhibited by increasing amounts of vanadate applied internally. The lack of effect of vanadate on the Na-Ca exchange is readily demonstrated in the experiment shown in Fig. 2. In this case the axon was dialysed with a solution free of ATP and containing 0.7 μM Ca²⁺ to maximise the Na-Ca exchange². The addition of 1 mM vanadate did not significantly change the levels of Ca efflux. (The small reduction seen could be due to inhibition of an uncoupled Ca efflux caused by ATP remaining inside the axon.) In addition, the removal of Ca_o and Na_o reversibly inhibited Ca efflux, bringing it to leakage levels. Both the absolute magnitude of the Na_o-Ca_o-activated Ca efflux and the leakage levels are similar to those found in axons not poisoned with vanadate, dialysed with the same [Ca²⁺]_i. Results similar to those shown in Figs 1 and 2 were seen in three other experiments. Figure 3 is a dose-response curve for the vanadate inhibition of the ATP-dependent uncoupled Ca efflux in axons dialysed with 0.05 μM Ca²⁺. From the curve, the K_i for vanadate is about 7 × 10⁻⁶ M. The inhibition by vanadate was always reversible.

If one accepts the possibility that the Ca pump is widespread in tissues and species⁷, the vanadate inhibition of the Ca-ATPase activity in red cells and the ATP-dependent Ca efflux in squid axons could be taken as a strong suggestion that they are due to an identical process.

It is intriguing that whereas ouabain selectively inhibits the Na-K pump, vanadate inhibits both the Na-K and the Ca pumps. Vanadate almost certainly inhibits the Na-K pump by slowing the conformational change that precedes the release of K from the enzyme-potassium complex, E₂(K), formed after dephosphorylation (refs 8, 9 and L. B. Karlsh and I. M. Glynn, unpublished); in addition, external K is required for inhibition⁸⁻¹⁰. There is no indication as to the mechanism for vanadate inhibition of the Ca-ATPase, although Na⁺ and Mg²⁺ are known to facilitate this inhibition⁵. The present experiments show that, similar to its effect on (Na⁺ + K⁺)ATPase, vanadate inhibits the Ca pump by acting at the intracellular surface. On the other hand, this inhibition was not dependent on external K (data not shown). Experiments are in progress to determine the sidedness of vanadate and cation interactions in the inhibition of the Ca pump in squid axons.

We acknowledge Professor R. L. Post for suggesting the use of vanadate. This work was supported by CONOCIT grant 51-0602.

Received 30 April; accepted 11 June 1979.

1. Baker, P. F. *Prog. Biophys. molec. Biol.* **24**, 177-223 (1972).
2. DiPolo, R. & Beaugé, L. A. *Nature* **278**, 271-273 (1979).
3. Beaugé, L. A. & Glynn, I. M. *Nature* **268**, 355-356 (1977).
4. Cantley, L. C. *et al. J. biol. Chem.* **252**, 7421-7422 (1977).
5. Bond, G. H. & Hudgins, P. M. *Fedn Proc.* **37**, 315 Abstr. (1978).
6. Brinley, F. J. & Mullins, L. J. *J. gen. Physiol.* **52**, 181-211 (1968).
7. Lew, V. L. *Nature* **274**, 421-422 (1978).
8. Beaugé, L. A. & DiPolo, R. *Biochim. biophys. Acta* **551**, 220-223 (1979).
9. Beaugé, L. A. in *2nd Conf. on Properties and Functions of Na, K-ATPase* (ed. Skou, J. C.) (Academic, New York, in the press).
10. Beaugé, L. A. & Glynn, I. M. *Nature* **272**, 551-552 (1978).

A selectively driven molecular clock

Daniel Hartl & Daniel Dykhuizen

Purdue University, Department of Biological Sciences,
West Lafayette, Indiana 47907

The problem of whether mutations that increase a species' fitness are incorporated into the genome as a function of generation interval or as a function of absolute time is of some interest because of the clock-like nature of amino acid substitutions known to have occurred in the evolution of certain well studied proteins. The molecular clock features a rate of amino acid substitution that is relatively constant when averaged over a sufficiently long time, and the constant rate of substitution is independent of the generation interval of the species involved^{1,2}. Furthermore, the variance in the rate of substitution is roughly twice the mean². Some authors have argued that the very existence of a molecular clock implies that most amino acid substitutions must be selectively neutral or nearly neutral³. On the other hand, there is no compelling reason to assume that selectively driven gene substitutions would not also exhibit clock-like behaviour, in which case the rate of increase in fitness of a population should be a function of absolute time rather than generation interval. In this report, we have investigated the issue of rate of increase in fitness using initially isogenic strains of *Escherichia coli* that had been allowed to evolve independently in glucose-limited chemostats⁴ for 500 h. The nutrient supply was adjusted in such a manner that some of the strains maintained a doubling time of 2.5 h and others a doubling time of 5.0 h, and the relative fitnesses of the evolving strains were measured periodically. Our main conclusions are that (1) selectively driven gene substitutions detectable as increases in fitness do exhibit clocklike behaviour, (2) when averaged over a sufficiently long time, the rate of increase in fitness is a function of absolute time and independent of generation interval, and (3) except for an initial period related to adjustment to chemostat conditions, the variance in the rate of increase in fitness is very nearly twice the mean.

At each generation interval, two isogenic strains were maintained in separate chemostats. One strain was DD616, which is sensitive to bacteriophage T5 (T5^S) and which carries *rpsL*, a non-reverting streptomycin-resistance mutation used to verify that the chemostats remained free of contamination; the other strain was selected as a spontaneous mutation in DD616 that

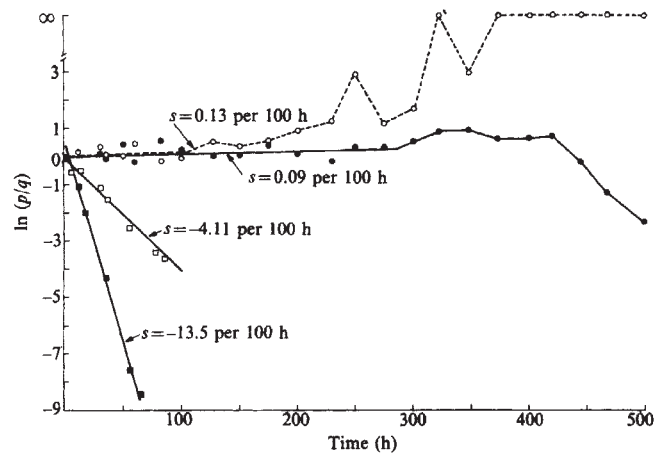


Fig. 1 Changes in the proportion of T5^R cells (p) relative to the proportion of T5^S cells (q) in chemostats. ●, Isogenic T5^R and T5^S cells maintained at a 2.5-h doubling time. ○, Isogenic T5^R and T5^S strains maintained at a 5.0-h doubling time. ■, Chemostat with 2.5-h doubling time containing unevolved T5^R strain against T5^S strain that had evolved for 200 h at a 5.0-h doubling time. □, Chemostat with 5.0-h doubling time showing genetic divergence between a T5^R strain grown for 300 h at a 2.5-h doubling time and a T5^S strain grown for 300 h at a 5.0-h doubling time.

confers resistance to T5 (T5^R), a phenotype associated with the *tonA* locus. Each chemostat contained 30 ml of medium and quickly reached an equilibrium density of about 5×10^8 cells per ml. Samples were withdrawn from the chemostats at 100-h intervals and aliquots containing approximately 10^{10} cells were frozen in 20% glycerol at -20°C for subsequent study.

The selective difference between any two strains was estimated by inoculating a glucose-limited chemostat with approximately 10^9 growing cells from each strain. The chemostat was maintained at either a 2.5-h or a 5.0-h generation time for 80–100 h, and at intervals the chemostat was sampled and the relative proportions of the two strains ascertained. The genetic marker used to distinguish the strains was resistance or sensitivity to T5, and chemostat samples were either plated on nonselective medium to estimate the total number of cells or were plated in the presence of an excess of T5 to estimate the number of T5^R cells; the ratio of T5^R cells to total cells estimates the proportion, p , of T5^R cells in the chemostat.

Selection in the chemostat is indicated by a systematic increase or decrease in the frequency of T5^R. Indeed, for a haploid population undergoing continuous growth⁵, $dp/dt = pqs$, where $q = 1 - p$ represents the proportion of T5^S cells at

Table 1 Fitness differences observed per 100 h for strains that had evolved for various lengths of time in either 2.5-h or 5.0-h chemostats

2.5-h chemostats		T5 ^R					
Time (h)		0	100	200	300	400	500
T5 ^S	0	0.09 ± 0.06					
	100	-12.82 ± 1.14*	-0.25 ± 0.16				
	200	-15.66 ± 2.83*	-0.44 ± 0.15†	0.34 ± 0.22	1.02 ± 0.28†		
	300		-1.16 ± 0.20*	-0.28 ± 0.21	0.22 ± 0.21	3.48 ± 0.37*	
	400		-5.48 ± 0.24*		-3.89 ± 0.18*	-1.55 ± 0.28*	-0.01 ± 0.20
	500		-5.98 ± 0.20*			-1.90 ± 0.30*	-0.71 ± 0.31
5.0-h chemostats		T5 ^R					
Time (h)		0	100	200	300	400	500
T5 ^S	0	0.13 ± 0.23					
	100	-9.63 ± 0.65*	-2.11 ± 0.27*				
	200	-12.90 ± 0.33*	-9.81 ± 0.52*	-0.07 ± 0.24	1.47 ± 0.70		
	300		-11.91 ± 0.21*	-3.60 ± 0.26*	-1.70 ± 0.13*	0.27 ± 0.11	
	400		-10.10 ± 0.40*		-2.31 ± 0.31*	-1.10 ± 0.17*	-0.23 ± 0.30
	500		-10.94 ± 0.29*			-0.88 ± 0.15*	-1.42 ± 0.17*

In each case the selection coefficient is that of T5^R relative to T5^S. Underlined values are means of two independent estimates that were not significantly different. All selection coefficients are estimated by the slope of a regression line, as shown in Fig. 1. * and † indicate values significantly different from 0 at 1% and 5% levels, respectively.

time t , and s is the selection coefficient denoting that e^s is the rate of increase in the number of $T5^R$ cells relative to a value of 1 for the rate of increase of $T5^S$. Thus, $\ln(p_i/q_i) = \ln(p_0/q_0) + st$, so a plot of $\ln(p/q)$ against t is linear with slope s .

Figure 1 shows the selective equivalence of the original, unevolved $T5^R$ and $T5^S$ strains when grown in chemostats at 2.5 h (filled circles) or 5.0 h (open circles). These control chemostats were established at the same time as the experimental ones and were identical to them, except that the controls were inoculated with equal numbers of $T5^R$ and $T5^S$, and the frequency of $T5^R$ was periodically estimated. In the 2.5-h chemostat the frequency of $T5^R$ remained at very nearly 50% for 275 h; the slope of the linear regression from 0 to 275 h is 0.09/100 h and is not significantly different from 0. After 275 h the frequency of $T5^R$ begins to change somewhat erratically, a phenomenon caused by periodic selection⁶—the chance occurrence of rare, favourable mutations that carry the associated $T5$ marker with them as they change in frequency. The onset of periodic selection in the 5.0-h chemostat is earlier (125 h) and more pronounced, but the initial slope of the regression (0.13/100 h for 0–100 h) is again not significantly different from 0.

Results of the experiments are summarised in Table 1, which presents the selection coefficient in favour of $T5^R$ in competition experiments involving $T5^R$ and $T5^S$ strains that had evolved independently for various lengths of time. The upper and lower matrices refer, respectively, to strains evolved and tested at a 2.5-h and a 5.0-h generation interval. In these matrices the difference between any number and the number immediately below it represents the increment in fitness achieved in the corresponding 100 h of evolution; 100-h increments in fitness are also given by the difference between any number and the number immediately to its left.

Fitness increments calculated from Table 1 are presented in Table 2. Those increments which did not differ significantly from 0 were entered as 0 in Table 2. (The two negative entries in Table 2 are probably spurious, as the probability of obtaining one or more spurious significances at the 1% or 5% level with 44 tests is 36% and 90%, respectively.) As indicated in Table 2,

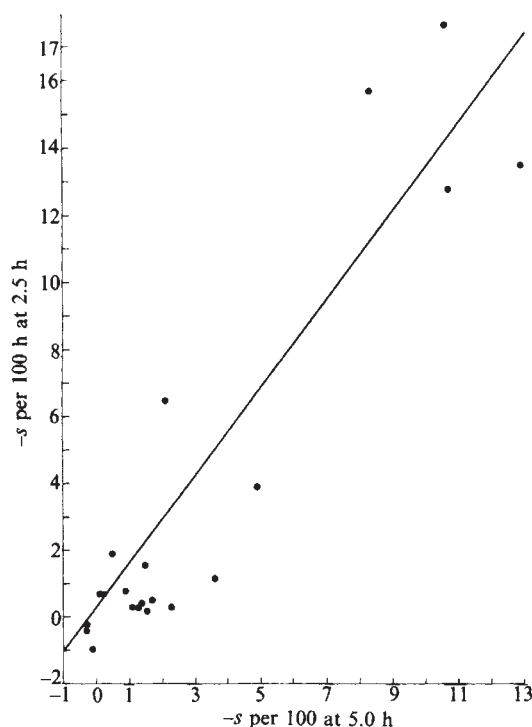


Fig. 2 Relationship between the observed fitness difference of two strains at a 2.5-h doubling time and the observed fitness difference of the identical strains at a 5.0-h doubling time. The regression is $y = 0.29 + 1.35x$.

Table 2 Fitness increments during various intervals of evolution in chemostats having either a 2.5-h(a) or a 5.0-h(b) doubling time

a	0–100	100–200	200–300	300–400	400–500
	12.91	0	0.72	4.32	0
	12.57	0	0	4.11	0
	15.22	0.78	0.80	5.03	0
		0.88	0	3.26	1.54
			0	2.34	1.19
	13.57	0.42	0.30	3.81	0.55
b					
	9.76	3.27	2.10	–1.81	0
	7.52	7.70	3.53	0	0
	3.09	9.74	3.17	1.37	1.19
		8.31	1.54	1.97	0.87
			1.90	1.21	–0.54
	6.79	7.26	2.45	0.55	0.30

Values were obtained from Table 1 by subtracting an entry from either the one immediately above it or the one immediately to its right. Differences not significantly different from 0 are entered as 0. Values beneath each column are unweighted means.

considerable variation in the fitness increments occurs. The large values obtained for 0–200 h undoubtedly arise from the selection of favourable mutations present in the initial inoculum of 2×10^9 cells. The somewhat more regular behaviour of the fitness increments in the period 200–500 h represents long-term adjustment to the chemostat based on incorporation of newly arising mutations as evidenced by periodic selection in the controls. The overall average fitness increment per 100 h is 2.99 in the 2.5-h chemostat and 3.00 in the 5.0-h chemostat. In the 200–500-h period, the average fitness increments per 100 h are 1.55 and 1.10 in the 2.5-h and 5.0-h chemostats, respectively; these values are not significantly different.

Thus, when averaged over a sufficiently long time, the rate of increase in fitness under selection is constant and independent of generation time. However, the comparisons in Table 2 would be invalid if a given fitness increment at 2.5 h were magnified to twice its value at 5.0 h. In the worst case the regression of fitness increment at 2.5 h on fitness increment at 5.0 h would have a slope of 0.5, for then the values in Table 2a would have to be corrected by multiplying by a factor of 2, and the average fitness increments would then be a function of generation interval.

Figure 2 shows 21 cases where two strains were competed at both 2.5 and 5.0 h. The regression is $y = 0.29 + 1.35x$, where y and x represent the selection coefficients at 2.5 and 5.0 h, respectively; the linear regression on x accounts for 85% of the variation in y . The intercept, 0.29, has a standard error of 0.67 and does not differ significantly from 0, as expected. The slope, 1.35 (with s.e.m. 0.14), differs significantly from 0.5 but its 98% confidence interval includes 1. The comparisons in Table 2 that imply a selectively driven molecular clock are therefore justified.

As for the variance in fitness increments, attention should be restricted to the 200–500 h period to exclude the large effects in the first 200 h due to pre-existing mutations in the initial inoculum. For this period, the overall mean of the 30 increments in Table 2 is 1.33 per 100 h and the variance is 2.58. As the variance is very nearly twice the mean, the selectively driven fitness clock is remarkably similar to the molecular clock based on comparisons of amino acid sequences.

This work was supported by NIH grant GM24886.

Received 25 May; accepted 1 August 1979.

1. Sarich, V. & Wilson, A. *Science* **154**, 1563–1566 (1966); *Proc. natn. Acad. Sci. U.S.A.* **58**, 142–148 (1967).
2. Fitch, W. M. in *Molecular Evolution* (ed. Ayala, F. J.) 160–178 (Sinauer, Sunderland, Massachusetts, 1976).
3. Nei, M. *Molecular Population Genetics and Evolution* (North-Holland, Amsterdam, 1975).
4. Dykhuizen, D. *Evolution* **32**, 125–150 (1978).
5. Crow, J. F. & Kimura, M. *An Introduction to Population Genetics Theory* (Harper and Row, New York, 1970).
6. Atwood, K. C., Schneider, L. K. & Ryan, F. J. *Cold Spring Harb. Symp. quant. Biol.* **16**, 345–355 (1951).

Macromolecular structural transitions in Pf1 filamentous bacterial virus

C. Nave, A. G. Fowler, S. Malsey, D. A. Marvin & H. Siegrist

European Molecular Biology Laboratory, Heidelberg, FRG

E. J. Wachtel

Department of Chemical Physics, Weizmann Institute of Science, Rehovot, Israel

The filamentous bacterial virus Pf1 is a simple model for biological filaments. We have studied the structure of the virion and report here that the helix parameters of Pf1 change sharply with temperature at about 8 °C. Local interactions between protein subunits change by only a few tenths of an ångström, but the changes are amplified between one end and the other of the virion to a rotation of 15 turns and a translation of 1,000 Å. The limited nature of the phase transition is probably due to the constraints of 'knobs-into-holes' interaction between side chains of adjacent α -helical protein subunits. Treatment of the virion with ether causes a rearrangement of protein subunits into sheets, with the α -helices normal to the plane of the sheet. This phase transition suggests a model for virion assembly in the bacterial membrane.

The Pf1 virion is a helical array of 46-residue α -helical protein subunits. It is 20,000 Å long by 60 Å in diameter, encapsulating a single-stranded circular DNA molecule at its core. Two types of X-ray fibre diffraction patterns from Pf1 have been obtained¹, both showing the same general distribution of intensity but each showing slightly different layer line positions (Fig. 1). Fibres prepared at room temperature show a pattern of layer lines indicating 4.40 units in one turn of the basic helix pitch, whereas fibres prepared at 4 °C show a pattern of layer lines indicating 4.46 units per turn (u/t). Although the local difference between the surface lattices of the two structures is slight, the effect is

cumulative (Fig. 2), so that over the 20,000 Å length of the virion, the transition causes one end of the virion to rotate with respect to the other by about 15 turns, at the same time causing the length of the virion to change by about 1,000 Å. The same sort of change can be induced in a wet gel. When the gel is prepared at room temperature, it shows features typical of the 4.40 u/t diffraction pattern; when the temperature is reduced, features typical of the 4.46 u/t pattern appear within the minimum time necessary to obtain a diffraction pattern. The transition in the gel is reversible. We have varied many parameters other than temperature, including hydration of the fibre, pH, ionic strength, nature of monovalent or divalent cations present, and organic solvents, but so far we have found no other treatment that causes this transition at room temperature.

It is convenient to discuss this transition with regard to a specific molecular model. The simplest model of the protein subunit that explains the diffraction data is a gently curved α -helix rod about 70 Å long by 10 Å in diameter with its long

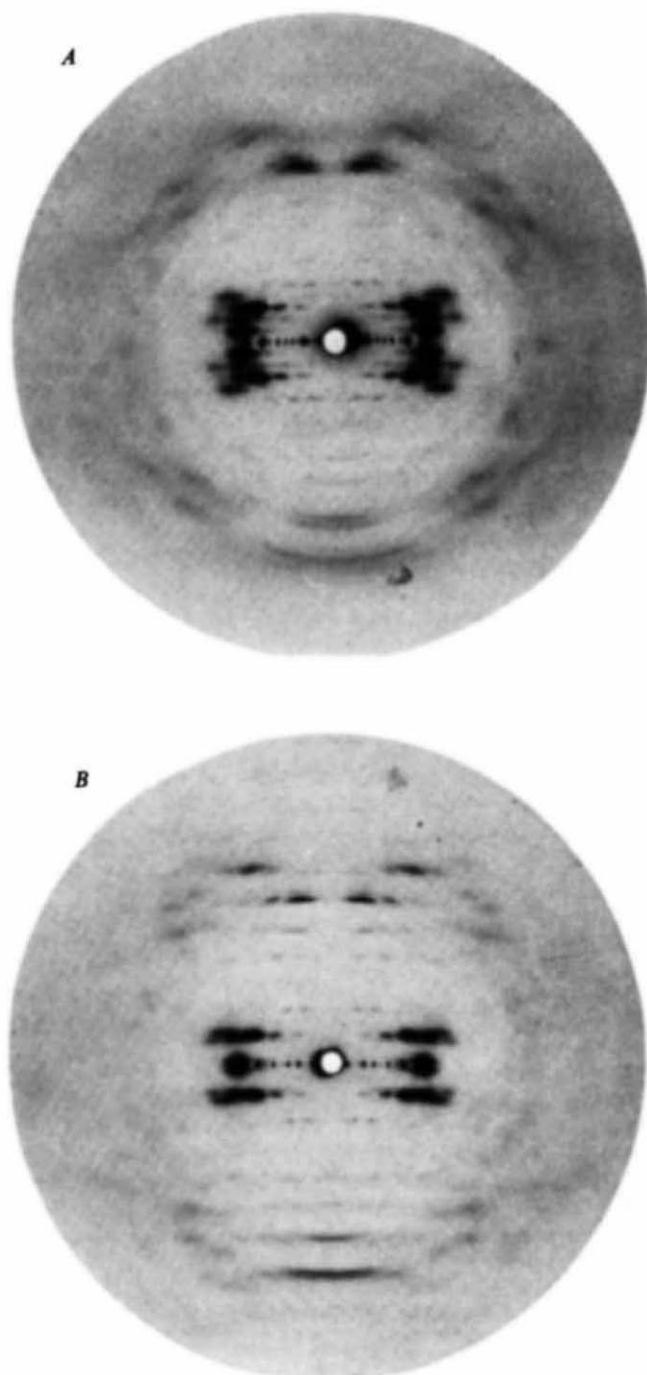


Fig. 1 X-ray diffraction patterns of Pf1 bacterial virus. X-ray fibre diffraction techniques have been described²⁰. Fibres were prepared at the temperatures indicated, but X-ray patterns were taken at room temperature after the conformation had been fixed in the dried fibre. A strong magnetic field (6–7 T) was used to improve orientation of the α -helical virions in the fibres, following X-ray diffraction experiments by Samulski²¹ and by J. Torbet in collaboration with us (report by Maret *et al.*²²) showing that α -helices and Pf1 orient parallel to a magnetic field. The fibre axis (and the virion axis) is vertical. The broad distribution of intensity on the pattern gives information about the subunit shape. Strong regions of intensity in the equatorial (horizontal) direction at 10 Å and in the meridional (vertical) direction at 5 Å indicate that most of the protein in the virion is in the form of α -helix with the α -helix roughly parallel to the virion axis. The horizontal lines of intensity (layer lines) give information about the virion symmetry. Near-meridional spacings on layer lines at orders of 0.06385 Å⁻¹ in the 4.40 u/t form and 0.06020 Å⁻¹ in the 4.46 u/t form give the pitch of the basic helix as 15.66 and 16.61 Å, respectively. The positions of the other layer lines give information about the number of equivalent protein subunits in one turn of the basic helix. These other layer lines change position relative to the basic set as a result of the temperature-induced transition. Because of an inherent ambiguity in fibre diffraction patterns, it is not clear^{3,4} whether the number of units per turn at room temperature is 4.40 or 5.40, but the detailed interaction between neighbouring subunits is virtually the same for the two alternatives, so this ambiguity does not affect the considerations discussed here. We choose 4.40 u/t for the purpose of discussion. **A**, 4.40 u/t , unit rise 3.56 Å, $a = 62.5$ Å, 92% relative humidity, specimen–film distance 7.50 cm, tilt 8°, disorientation half-width 2.3°. Fibre AF 12/2, prepared at room temperature, film AF 055, $\times 0.98$. The layer line splitting in the 5 Å meridional region can be approximately explained by a left-hand coiled-coil⁵ of radius about 1 Å and pitch 150 Å. **B**, 4.46 u/t , unit rise 3.73 Å, $a = 61.9$ Å, 98% relative humidity, specimen–film distance 8.28 cm, tilt 12°, disorientation half-width 1.5°. Fibre AF 35/1, prepared at 4 °C, film 4158, $\times 0.9$.

axis oriented roughly parallel to the virion axis and also sloping radially². Variants of this model have also been discussed³. Each subunit is in contact with neighbours 5 units and 9 units along the basic virus helix (Fig. 2). These contacts can be considered by imagining the axes of the α -helices to move smoothly along the 5-start or 9-start directions, so as to sweep out sets of twisted ribbons, or helicoids^{2,4}.

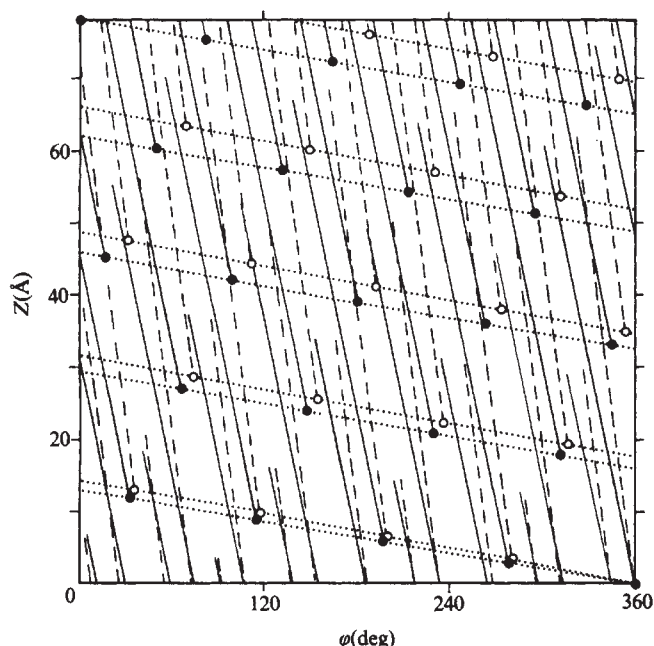


Fig. 2 Surface lattices of Pf1 helices with 4.40 u/t (filled circles and solid lines) and 4.46 u/t (open circles and dashed lines). The dotted lines indicate the basic helices. The surface lattice is generated by projecting equivalent points along radii onto a cylinder of 11.4 Å radius, and then cutting the cylinder vertically to give a two-dimensional representation. The circles represent the projections of equivalent points and the lines represent the projection of the α -helix axes.

Details of the model become less important if we discuss the changes in the structure during the transition rather than the absolute structures. The axis of the subunit in the 4.46 u/t model is more nearly parallel to the axis of the virion than in the 4.40 u/t model (Fig. 2), but the shape of the protein subunits and the interactions between them remain virtually unchanged during the transition. The close relationship between the two structures can be seen by comparing the diffraction data from the strong J_5 and J_9 layer lines of the two structures (Fig. 3A and B). Although the ζ (vertical) positions of these layer lines change as a result of the transition, because of the change in helix symmetry, the intensity distributions along these layer lines (which reflect the molecular structure of the subunit) show virtually no change. The similarity between amplitudes predicted for the two models (Fig. 3C and D) is comparable to the similarity between amplitudes observed in the diffraction experiment. Large changes within the subunit, such as bending at a hinge, are not supported by the data. Very small and subtle changes in neighbour-neighbour interactions are sufficient to explain the observations. In particular, the fit of side chains of one subunit into the space between side chains of its neighbours ('knobs-into-holes')⁵ is unchanged. This similarity was defined more precisely by calculating distances between pairs of β -carbons on neighbouring subunits where the β -carbons approach one another to within 5 Å. The specific pairs of atoms involved in such contacts do not alter between the 4.40 u/t and the 4.46 u/t models, and even the absolute distances between atoms in these pairs change by only a few tenths of an ångström.

The transition appears to be discontinuous and limited. That is, no intermediate forms between 4.40 u/t and 4.46 u/t have been observed, and no other forms have been observed outside

these two extremes. The limitation follows directly from consideration of the knobs-into-holes packing in the model. Knobs-into-holes interactions constrain the crossing angle between the axes of neighbouring α -helices⁵. Ideal α -helices with an internal rotation per residue of 100° will cross at an angle of 20°. If the crossing angle changes from 20°, the internal rotation per residue must change correspondingly. However, stereochemical constraints limit the rotation per residue to within a degree or so of 100° (ref. 6), so crossing angles of roughly parallel helices are near 20° (ref. 7). Packing between neighbouring α -helical subunits in the Pf1 model is also constrained by knobs-into-holes considerations². In order to retain unchanged knobs-into-

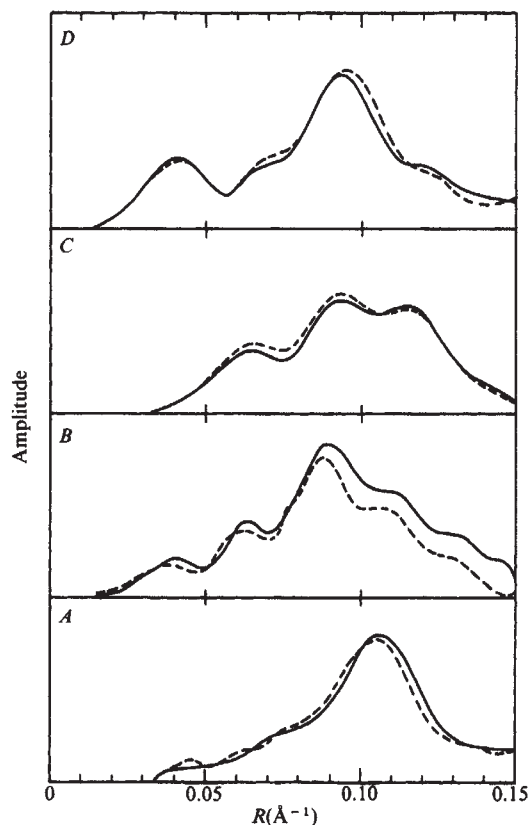


Fig. 3 Distribution of diffraction amplitude along layer lines in the strong 10 Å near-equatorial region. Solid curves represent 4.40 u/t amplitudes and dashed curves represent 4.46 u/t amplitudes. Intensity measurements and Fourier transform calculations have been described². The models have side chains attached in the most probable conformation to the α -helix backbone². The two strongest contributions of intensity in the 10 Å region are due to the periodicities of the 5-start and the 9-start helicoids, which are expressed in the helical Fourier transform as the J_5 and J_9 Bessel functions. The agreement between calculated and observed amplitudes over the whole diffraction pattern is similar to that of previous models². The remaining differences between calculated and observed amplitudes are of the order usually found in fitting models to fibre diffraction data, and do not affect our discussion. A, Observed amplitude for the J_9 layer line. 4.40 u/t data for the layer line at $\zeta = 0.0128 \text{ Å}^{-1}$ (fibre AF 12/2, film 4055) and 4.46 u/t data for the layer line at $\zeta = 0.0046 \text{ Å}^{-1}$ (fibre AF 35/1, film 4158). B, Observed amplitude for the J_5 layer line. 4.40 u/t data for the layer line at $\zeta = 0.0383 \text{ Å}^{-1}$ and 4.46 u/t data for the layer line at $\zeta = 0.0278 \text{ Å}^{-1}$. Film numbers as in A. The apparent discrepancy between the two curves beyond 0.1 Å^{-1} is due in part to the contribution of the J_{17} Bessel function, which cannot be resolved from J_5 in the 4.40 u/t data but is resolved (and not included) on the 4.46 u/t data. C, Calculated amplitude for the J_5 layer line (ζ as in A). D, Calculated amplitude for the J_9 layer line (ζ as in B).

holes packing during the transition from a 4.40 u/t to a 4.46 u/t model, the rotation per residue within the α -helix must be changed by about 0.5°. The absence of virions with less than 4.40 u/t or greater than 4.46 u/t can then be explained as a result of a limitation on internal twist within the α -helix that is coupled by knobs-into-holes packing to the number of units per turn in the virion. This same constraint was shown in a different

way using a non-convergent energy minimisation technique which gave a local minimum for a series of models in the region of the transition¹.

Stereochemical constraints satisfactorily explain the limited nature of the transition, but they do not directly explain why the transition is driven by temperature. It is apparent that the stability of the 4.46 *u/t* helix relative to the stability of the 4.40 *u/t* helix changes with temperature. This effect could be due to temperature-dependent terms in the molecular interaction potentials controlling contact distances or patterns of attraction and repulsion between side chains. Changes in the interaction potentials could alter either the internal twist of the α -helix or the details of knobs-into-holes packing.

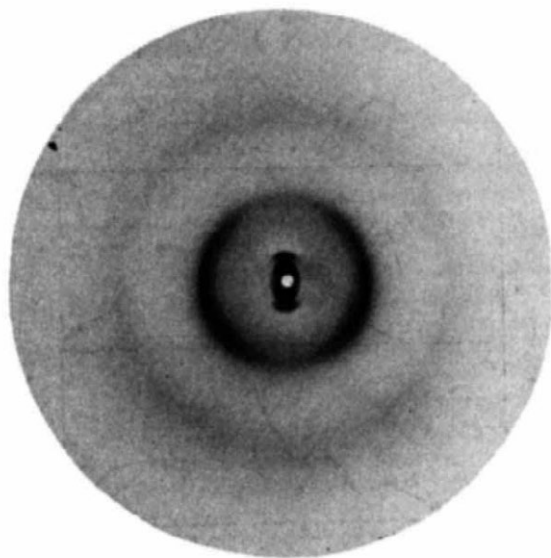


Fig. 4 X-ray diffraction pattern of ether-induced sheets of Pf1 bacterial virus. Plane of the sheet is horizontal. Sheets were prepared by shaking a purified solution of Pf1 at 5 mg ml⁻¹ for 120 min at room temperature with an equal volume of ether⁴. The sheets were centrifuged at 30,000 r.p.m. for 3 h at 15 °C in the Beckman SW 50.1 rotor, and resuspended at a concentration of 25 mg ml⁻¹ in 0.005M Tris-HCl, pH 8.1. A 0.5 ml drop of the concentrated solution was placed on a silicone-treated glass slide²³ and allowed to dry at 100% relative humidity, 4 °C, for 160 h in a 7-T magnetic field, with the plane of the slide perpendicular to the magnetic lines of force. A thin slice of the drop was cut free from the slide and mounted in the X-ray diffraction camera with the X-ray beam parallel to the plane defined by the slide. The sharp vertical reflections are at orders of 55.6 Å; on some patterns reflections up to the fifth order have been observed. The strong region of intensity in the horizontal direction is at 9 Å. Sample AF 33/2, film 4146, 75% relative humidity, specimen-film distance 7.4 cm, $\times 1.0$. The same sample at 98% relative humidity expanded slightly to give sharp vertical reflections at orders of 60.1 Å.

Solvent is likely to play an important part in such phenomena. Even internal structure may be affected by solvent. Hydration behaviour can be subdivided⁶ into ionic or electrostatic effects, hydrogen bonding of water to the solute, and hydrophobic effects. All of these are temperature dependent to some degree, and may play some part in the Pf1 transition. The electrostatic effect⁹ is an especially likely candidate for the structural transition in Pf1. Theoretical calculations show that the dielectric constant of the solvent has an important effect on the conformation of proteins¹⁰. The dielectric constant of water increases as the temperature is reduced¹¹. In the presence of glycerol (which has a dielectric constant about half that of water), the 4.40 *u/t* form is maintained even when fibres are made at 4 °C. This fact is consistent with the idea that the temperature-induced conformational change is due to the increase in the dielectric constant of the medium with decreasing temperature. Special cryosolvents¹² that have been developed

to counteract the effect of temperature on pH may be useful to study the structural transition in Pf1.

It is possible that the system is instead controlled by entropy effects, resulting from vibration or libration of the protein molecules. Calorimetry could distinguish between the two possible mechanisms.

More dramatic changes in Pf1 structure can be induced by brief treatment with ether. The virions swell to hollow cylinders about three times their original diameter but only one tenth the normal length¹³; further treatment leads to flat sheets of protein⁴. X-ray diffraction studies of oriented stacked sheets show a strong 9 Å region of intensity in the direction parallel to the plane of the sheet, and sharp reflections at orders of about 55–60 Å in the direction perpendicular to the plane of the sheet (Fig. 4). This kind of diffraction pattern is expected if the Pf1 α -helical coat protein subunits are oriented with their long axes roughly perpendicular to the plane of the sheet. If the 46-residue α -helices were oriented exactly perpendicular to the sheets, the sheets would be 70 Å thick, but if some of the neighbour-neighbour side chain interactions in the sheet are the same as in the 5-start helicoids of the virion, the α -helices will be slightly skew and the sheets will be only 50–60 Å thick. The 9 Å spacing in the plane of the sheet corresponds to the distance between adjacent α -helices. Patches of coat protein in the bacterial membrane with the α -helical protein spanning the 50 Å thick membrane have been proposed as precursors of virus assembly^{2,14–16}. These pure viral protein sheets are useful models for the study of assembly of the virion.

The Pf1 virion is an example of a system in which small local changes in molecular structure are amplified into large reversible changes in a molecular assembly. The changes within the α -helix subunit between the 4.40 *u/t* and the 4.46 *u/t* models are similar in magnitude to the calculated room temperature vibrations within an α -helix¹⁷. At the transition temperature, Pf1 will be delicately poised between the two structural forms, and little energy would be required to induce an oscillation between the two forms. This type of induced oscillation between two structural forms of an α -helical assembly could generate a travelling wave in other α -helical assemblies such as myosin or membrane pores^{18,19}.

We thank J. E. Ladner and R. C. Ladner for discussions and for help with computer programs; H. Scheraga and V. Yakhot for helpful discussions; and the Deutsches Krebsforschungszentrum, Heidelberg, for computing facilities.

Received 4 May; accepted 30 July 1979.

1. Wachtel, E. J., Marvin, F. J. & Marvin, D. A. *J. molec. Biol.* **107**, 379–383 (1976).
2. Marvin, D. A. & Wachtel, E. J. *Nature* **253**, 19–23 (1975); *Phil. Trans. R. Soc. B* **276**, 81–98 (1976).
3. Makowski, L. & Caspar, D. L. D. in *The Single-Stranded DNA Phages* (eds Denhardt, D. T., Dressler, D. & Ray, D. S.) 627–643 (Cold Spring Harbor Laboratory, New York, 1978).
4. Marvin, D. A. in *The Single-Stranded DNA Phages* (eds Denhardt, D. T., Dressler, D. & Ray, D. S.) 583–603 (Cold Spring Harbor Laboratory, New York, 1978).
5. Crick, F. H. C. *Acta crystallogr.* **6**, 685–689; 689–697 (1953).
6. Fraser, R. D. B. & MacRae, T. P. *Conformation in Fibrous Proteins* (Academic, New York, 1973).
7. Chothia, C., Levitt, M. & Richardson, D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4130–4134 (1977).
8. Franks, F. *Phil. Trans. R. Soc. B* **278**, 33–57 (1977).
9. Record, M. T. Jr., Anderson, C. F. & Lohman, T. M. *Q. Rev. Biophys.* **11**, 103–178 (1978).
10. Warshel, A. & Levitt, M. *J. molec. Biol.* **103**, 227–249 (1976).
11. Eisenberg, D. & Kauzmann, W. *The Structure and Properties of Water* (Oxford University Press, 1969).
12. Douzou, P. *Cryobiochemistry* (Academic, New York, 1977).
13. Amako, K. & Yasunaka, K. *Nature* **267**, 862–863 (1977).
14. Marvin, D. A. & Hohn, B. *Bact. Rev.* **33**, 172–209 (1969).
15. Wickner, W. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1159–1163 (1976).
16. Webster, R. E. & Cashman, J. S. in *The Single-Stranded DNA Phages* (eds Denhardt, D. T., Dressler, D. & Ray, D. S.) 557–569 (Cold Spring Harbor Laboratory, New York, 1978).
17. Petricolas, W. L. *Biopolymers* **18**, 747–755 (1979).
18. Marvin, D. A., Pigram, W. J., Wiseman, R. L., Wachtel, E. J. & Marvin, F. J. *J. molec. Biol.* **88**, 581–600 (1974).
19. Dunker, A. K. & Marvin, D. A. *J. theor. Biol.* **72**, 9–16 (1978).
20. Marvin, D. A., Wiseman, R. L. & Wachtel, E. J. *J. molec. Biol.* **82**, 121–138 (1974).
21. Samulski, E. T. & Tobolsky, A. V. *Biopolymers* **10**, 1013–1019 (1971).
22. Maret, G. *et al.* in *Nonlinear Behaviour of Molecules, Atoms and Ions in Electric, Magnetic or Electromagnetic Fields*, 477–483 (Elsevier, Amsterdam, 1979).
23. Burge, R. E., Fowler, A. G. & Reaveley, D. A. *J. molec. Biol.* **117**, 927–953 (1977).

Local destabilisation of a DNA double helix by a T-T wobble pair

A. G. Cornelis, Jeroen H. J. Haasnoot, Jan F. den Hartog, M. de Rooij, Jacques H. van Boom & Altona Cornelis

Gorlaeus Laboratories, State University, PO Box 9502, 2300 RA Leiden, The Netherlands

Nuclear magnetic resonance is a technique which permits direct observation of the Watson-Crick hydrogen-bonded ring imino protons (guanine N₁H and thymine N₃H). As the formation and disruption of hydrogen bonds of double-helical RNA and DNA structures are key events during various biological processes, NMR thus provides a useful tool for studying the fluctuational mobility of the individual base pairs. Indeed, several NMR studies of oligo- and polynucleotides have been carried out to probe the structure and dynamics of nucleic acids in solution (for a review see ref. 1). The present study constitutes the first part of our attempt to assess the influence of non-complementary base pairs on the stability of nucleic acid double helices. We report the spectral assignment and temperature-dependent NMR profiles of the hydrogen-bonded imino protons of the two DNA fragments shown in Fig. 1. The assignment is based solely on experimental grounds using the principle of chemical modification. It will be demonstrated that the introduction of a non-complementary (wobble) base pair in a DNA duplex introduces an extra melting site in addition to the sequential melting which starts with the terminal base pairs in the double helix structure.

The oligodeoxynucleotides were synthesised using an improved phosphotriester approach². Full details of preparation and purification will be published elsewhere³. Figure 2 presents the high-resolution 360 MHz ¹H-NMR spectrum between 12.0 and 14.5 p.p.m. of fragment 12 in H₂O as a function of temperature. Sequential linewidth and intensity changes with increasing temperature allow assignment of all signals to specific NH-Watson-Crick hydrogen-bonded protons⁴. At low temperatures (−3.5 °C) six resonances, two of which coincide at ~13.3 p.p.m., are observed. On raising the temperature the intensity of the signal at 13.3 p.p.m. falls from two to one proton, identifying that the faded signal as the fraying terminal A·T base pair 1. Between 10 and 30 °C the resonance at 13.8 p.p.m. gradually disappears and is therefore assigned to A·T base pair 2. The signal at 12.5 p.p.m. is next to follow (G·C base pair 3).



Fig. 1 Base sequence of the self-complementary deoxy-oligonucleotides studied. Base pair numbering as shown; they are equivalent pairwise because of the two-fold symmetry of the double helix.

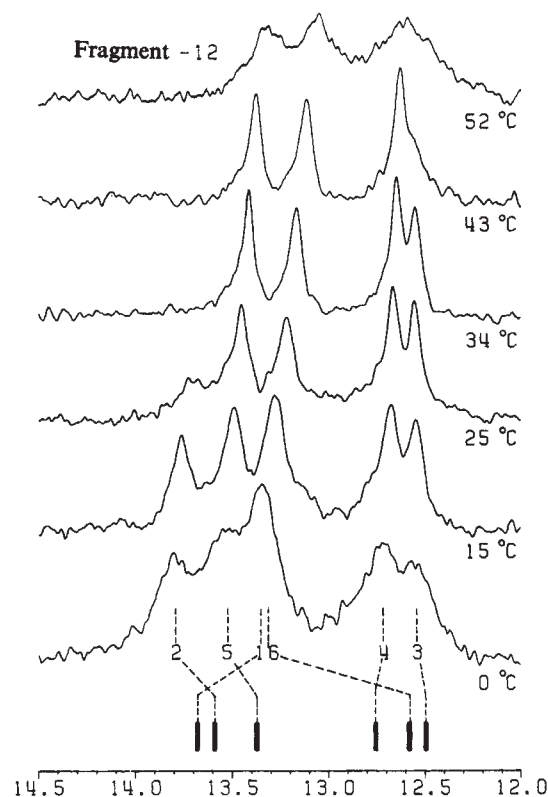


Fig. 2 Low-field region of the 360 MHz ¹H-NMR spectra of fragment 12 at various temperatures. Only a representative (6 out of 15) set of spectra is shown. The solid lines (bottom) represent the calculated shifts (see text) for the hydrogen-bonded imino protons in a B-DNA helix. The NMR samples were prepared by dissolving ~3 mg of the oligodeoxynucleotide in 270 µl of a buffer solution containing 0.5 M NaCl, 0.1 M sodium cacodylate (pH 7.0) and 5% v/v D₂O. High resolution proton NMR spectra were recorded on a Bruker HX-360 NMR spectrometer interfaced with a BNC-12 computer system and equipped with a standard temperature unit, using micro NMR sample tubes with a cylindrical cavity. The spectrometer was operated in the correlation mode¹⁴, using an internal deuterium field/frequency lock. Chemical shifts were measured relative to the solvent H₂O peak and converted to DSS reference by correcting for the H₂O to DDS chemical shift using the appropriate temperature calibration curve. All reported shifts are denoted using the δ-convention, that is, positive values represent shifts to lower field.

At 52 °C only three signals remain. As G·C base pairs resonate at ~1 p.p.m. to higher field than corresponding A·T base pairs⁵, the signal at 12.7 p.p.m. is assigned to the G·C base pair 4. Increasing the temperature to 56 °C results in the disappearance of all resonances indicating that the breaking of the last six base pairs is an 'all or none' process. The two A·T base pairs 5 and 6 can be assigned unambiguously by comparing the low temperature spectra of fragment 12 and fragment 13 (Fig. 3). In the spectrum of fragment 13 five out of six signals in the range of 12–14 p.p.m. display virtually the same chemical shifts as found in fragment 12; it is therefore concluded that the two fragments have similar conformations for the base pair sequence 1–5 and that the same assignment holds for fragment 13. The important difference, however, between the spectra of the two fragments is a large downfield shift of one of the resonances at 13.3 p.p.m. in fragment 12 to 14.3 p.p.m. in fragment 13. This shift is obviously due to the effect of separating base pairs 6 and 6' by insertion of a T-T pair. The insertion moves the imino proton of base pair 6 away from the shielding cone of the A·T base pair 6' and vice versa. Therefore, the resonance at 14.3 p.p.m. (and one of the resonances at 13.3 p.p.m. in fragment 12) is assigned to base pair

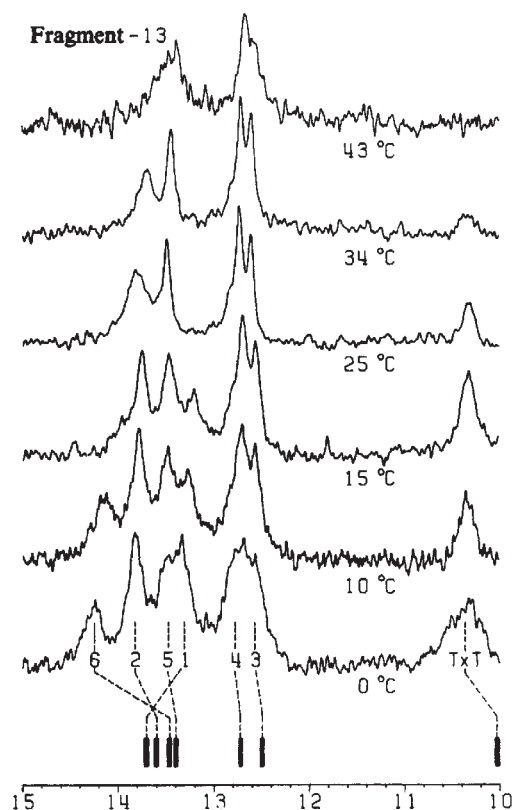


Fig. 3 Low-field region of the 360 MHz ^1H -NMR spectra of fragment 13 at representative temperatures. For further details see Fig. 2 legend.

6, and thus the remaining signal at 13.6 p.p.m. can be ascribed to base pair 5. An extra signal is displayed by fragment 13 at relatively high field (10.3 p.p.m.), which is assigned to the non-hydrogen-bonded imino protons of the thymidyl residues in the middle of the strands. The temperature profile of fragment 13 is consistent with this assignment, but shows two 'melting sites' instead of one (see below).

It is now generally recognised that oligodeoxynucleotides tend to take up a B-DNA-like conformation in aqueous solution⁶⁻⁸. The computer program CARSHI (written in FORTRAN IV, using the parametrisation of Giessner-Prettre and Pullman⁹) was used to calculate the ring current shielding due to the nearest and next-nearest neighbouring base pairs in a B-DNA double-helical structure¹⁰. These values were added to the estimate chemical shift for 'isolated' Watson-Crick base pairs⁵ ($\delta_{\text{G-C}} = 13.6$ p.p.m.; $\delta_{\text{A-T}} = 14.6$ p.p.m.), yielding the calculated shifts shown as solid lines in Figs 2 and 3. To determine the chemical shift of the 'free' (non-hydrogen-bonded) imino proton of a thymine residue the spectra of thymidine 5'-O-methyl phosphate and of d(TTTT) were recorded at 0 °C (~3 mM, pH = 7). Both compounds displayed a very broad resonance at 11.2 p.p.m., which sharpened when the exchange rate was suppressed by lowering either the temperature (to -9 °C) or the pH (to 4.0).

The agreement between calculated and observed chemical shifts seems reasonable, taking into account the possible and unknown effects of residual fraying and/or small deviations from the accepted B-DNA structure. However, it is clear that assignment of resonances close to one another on the basis of ring current shift calculations must be regarded with caution.

Fragment 12 represents the first oligodeoxynucleotide studied by nuclear magnetic resonance in which A-T base pairs in the middle of the strands are flanked on both sides by G-C base pairs. A-T-rich regions are known¹¹ to be thermally less stable than G-C-rich regions in long DNA fragments. One could

therefore expect that the central A-T region in fragment 12 would 'melt out' before the G-C pairs disappeared. This is not the case; our experiments clearly indicate that all four central A-T base pairs in fragment 12 remain intact until the last G-C base pairs 4 and 4' are disrupted. Therefore, we conclude that a 'G-C protected' sequence of four A-T pairs is too short for such a local destabilisation of the duplex.

For fragment 13, it is certainly important to determine whether the T-T pair loops out from the helix or whether it forms a wobble pair. The present experiments give the following pertinent information: the chemical shift of the T-T imino protons are ~1 p.p.m. to higher field than that observed for the 'free' imino proton of a thymine residue (see above). This large value for the upfield shift requires that this 'free' imino proton is positioned in the shielding cone of the neighbouring adenine base. Therefore, we propose that fragment 13 has a marked tendency to form a continuous, regular B-DNA-like structure with a T-T wobble pair embedded in the middle of the helix. Shift calculations agree well with this proposition (Fig. 3). Interestingly, a similar occurrence was recently reported in a temperature-jump study¹² of anticodon-anticodon interactions between various tRNAs. It was found that the U-U wobble in the middle position of the interacting codons behaves in a unique way in the sense that it apparently does not oppose the formation of stable complexes, whereas the other mismatches tested (C-C, A-A, G-G, G-A, C-A) seem to prohibit complex formation. Thus, the continuity of the DNA double helix with a T-T wobble pair may be found to be an exception rather than a rule, but further work is necessary to clarify this point.

The temperature profile of fragment 13 displays several interesting features. Compared with fragment 12, the NMR melting temperature is lowered by ~11 °C, indicating that the insertion of the T-T wobble pair significantly lowers the overall stability of the double helix. This observation agrees with thermal elution studies of oligodeoxyadenylates on cellulose-pdTp. The replacement of an internal dA- by a dT-residue lowered the elution temperature (which is compatible with the T_m obtained from solution thermal denaturation studies) by about 15 °C (ref. 13). An explanation of this phenomenon may be inferred from Fig. 3. The resonances of the T-T wobble pair and A-T base pair 6 located in the middle of the double-stranded structure vanish within the same temperature range as the terminal A-T base-pair 1. This implies that the wobble introduces an extra melting site in the double-helical structure. Thus, the breaking of the Watson-Crick hydrogen bonds occurs on both sides of the two G-C pairs, causing fragment 13 to be destabilised relative to fragment 12.

This study clearly demonstrates that high-field nuclear magnetic resonance techniques combined with 'tailor-made' oligonucleotides significantly increases our understanding of the structural features and dynamics of this type of important biological model compound.

This research was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO). We thank Mr K. Dijkstra and Dr R. Kaptein for technical assistance. See ref. 15 for the previous paper in this series.

Received 28 February; accepted 30 July 1979.

- Hilbers, C. W. in *Magnetic Resonance Studies in Biology* (ed. Shulman, R. G.) (in the press).
- Boom, J. H. van, Burgers, P. M. J. & Deursen, P. H. van *Tetrahedron Lett.* 869-872 (1976).
- Rooij, J. F. M. de, Wille-Hazeleger, G., Deursen, P. H. van, Serdy, J. & Boom, J. H. van *Recl. Trav. Chim. Pays-Bas* (in the press).
- Patel, D. J. *Biochemistry* 13, 2396-2402 (1974).
- Patel, D. J. & Tonelli, A. E. *Biopolymers* 13, 1943-1964 (1974).
- Altona, C., Boom, J. H. van & Haasnoot, C. A. G. *Eur. J. Biochem.* 71, 557-562 (1976).
- Patel, D. J. *Biopolymers* 16, 1635-1656 (1977).
- Selsing, E., Wells, R. D., Early, T. A. & Kearns, D. R. *Nature* 275, 249-250 (1978).
- Giessner-Prettre, C. & Pullman, B. *J. theor. Biol.* 27, 87-95 (1970).
- Arnott, S. & Hukins, D. W. L. *Biochem. biophys. Res. Commun.* 47, 1504-1509 (1972).
- Lehninger, A. L. in *Biochemistry* 2nd edn, 876 (Worth, 1975).
- Grosjean, H. J., Henau, S. de, & Crothers, D. M. *Proc. natn. Acad. Sci. U.S.A.* 75, 610-614 (1978).
- Gillam, S., Waterman, K. & Smith, M. *Nucleic Acids Res.* 2, 625-634 (1975).
- Dadok, J. & Sprecher, R. F. *J. Magnetic Resonance* 13, 243-248 (1974).
- Haasnoot, C. A. G. & Altona, C. *Nucleic Acids Res.* 6, 1135-1149 (1979).

Thromboxane molecules do not adopt the prostaglandin hairpin conformation

D. A. Langs, S. Fortier, M. G. Erman & G. T. DeTitta

Medical Foundation of Buffalo, Inc., 73 High Street, Buffalo, New York 14203

The hairpin conformational hypothesis¹ has been proposed to rationalise much of the structure-activity and receptor-binding data which have accumulated for the prostaglandin (PG) hormones. The hairpin conformation, thought to be necessary for PG activity, requires that the α - and ω -chains of the molecule be extended and in parallel alignment, separated by a van der Waals contact distance for the full length of the chains, with the ends of the chains approximately 5.5 Å apart. The similarity between the structures of the thromboxanes (TXs) and the PGs suggests that the profile of activity of TXs, like that of PGs, centres on subtle conformational variation of the hairpin geometry. Thromboxane B₂ (TXB₂) is a stable hydrolysis product² of a highly reactive, short-lived intermediate³, thromboxane A₂ (TXA₂), which is formed from the prostaglandin endoperoxide (PGH₂) as indicated in Fig. 1. An examination of molecular models of TXA₂ and TXB₂ suggests that the structural differences between the ring moieties may have much less influence in altering the side-chain conformations of TXs than do substituents on the relatively more flexible cyclopentane ring of a PG molecule. We report here the first diffraction analysis of a thromboxane structure and note that the molecular conformation is not hairpin shaped.

Thromboxane B₂ was crystallised from ethyl acetate as colourless, transparent, monoclinic laths following suggestions offered by Hanessian⁴. The space group is P2₁, with unit cell data $a = 9.971(3)$, $b = 40.45(2)$, $c = 5.510(2)$ Å, $\beta = 102.5(1)^\circ$, $Z = 4$, D_{obs} (floatation) = 1.141 g cm^{-3} , D_{calc} = 1.138 g cm^{-3} . There are two crystallographically independent TXB₂ molecules in the unit cell. The diffraction intensities were recorded from a prismatic specimen ($0.12 \times 0.12 \times 0.40 \text{ mm}^3$) in the θ -2 θ scan mode on an automated diffractometer fitted with a CuK α source. A total of 2,272 independent reflections were measured out to a 2θ angle of 100° . Of these, only 1,227 intensities were considered observable ($I \geq 2\sigma I$). The crystal structure was determined by direct methods and refined isotropically by full matrix least-squares procedures to the current residual $R = 0.145$. The general weakness of the diffraction data is consistent with the high thermal motion observed for the terminal carbon atoms of the ω -chains and the disorder in O₍₁₁₎ (α and β epimers) for one of the molecules in the structure.

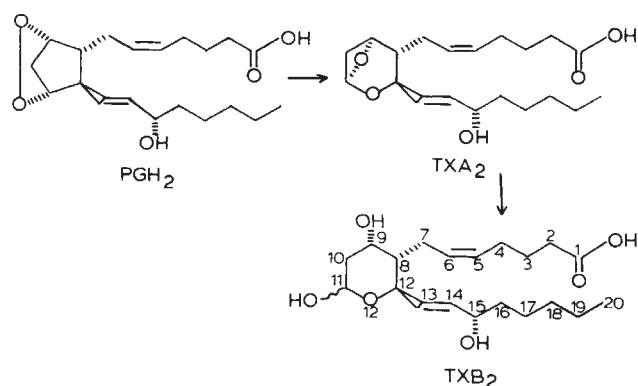


Fig. 1 Chemical formulae and numbering scheme for TXB₂ and its immediate precursors.

The two crystallographically independent molecules and their overlapped composite are shown in Fig. 2. The composite illustration was derived by overlapping corresponding atoms of both molecules in the least-squares sense⁵ for the common contiguous 12-atom moiety from C₍₆₎ to C₍₁₅₎, excluding O₍₁₁₎. Note the torsional dissimilarities at the C₍₄₎-C₍₅₎ and C₍₆₎-C₍₇₎ bonds (Table 1). The carboxyl side chains bend sharply towards either the α or β face of the molecule. For convenience, we refer to these as the α -tail and the β -tail 'scorpion' conformers, respectively. The TXB₂ ω -chain conformations are similar but distinct. The C₍₁₆₎-C₍₁₇₎ bonds are both +synclinal rather than transplanar, as noted for every one of 12 PG ω -chain conformations studied by diffraction methods. The kinking of the chain at the C₍₁₆₎-C₍₁₇₎ bond has the effect of foreshortening the ω -chain such that it extends only as far as C₍₄₎ for its terminal contact with the α -chain. It can be seen from the molecular overlap figure that the C₍₁₃₎-C₍₁₄₎-C₍₁₅₎-C₍₁₆₎ torsion angles, which define the orientations of the tails of the ω -chains and which differ by 16° , cause these tails to be slightly shifted either up or down to achieve this terminal contact with C₍₄₎. The end of the carboxyl chain is not stabilised by van der Waals association with the ω -chain as is the case in the hairpin conformation, and may be conformationally flexible with regard to the C₍₄₎-C₍₅₎ hinge point. Although molecular models might suggest that viable hairpin conformers of TXs may be formed close to the mean plane of the pyranose ring, we consider that such hairpin models generally crowd the side chains more closely together than if they had emanated from a cyclopentane head group. A +anticlinal C₍₆₎-C₍₇₎ bond conformation has previously been observed for each of four conformers of PGF_{2 α} (refs 6, 7). The

Table 1 Thromboxane B₂ torsion angles for β -tail and α -tail conformers

O _(1A)	C ₍₁₎	C ₍₂₎	C ₍₃₎	2(10)
C ₍₁₎	C ₍₂₎	C ₍₃₎	C ₍₄₎	-177(176)
C ₍₃₎	C ₍₄₎	C ₍₅₎	C ₍₆₎	-142(144)
C ₍₅₎	C ₍₆₎	C ₍₇₎	C ₍₈₎	91(-148)
C ₍₆₎	C ₍₇₎	C ₍₈₎	C ₍₁₂₎	173(-177)
C ₍₇₎	C ₍₈₎	C ₍₉₎	C ₍₁₀₎	176(175)
C ₍₇₎	C ₍₈₎	C ₍₁₂₎	O ₍₁₂₎	176(180)
C ₍₈₎	C ₍₁₂₎	C ₍₁₃₎	C ₍₁₄₎	-144(-123)
C ₍₉₎	C ₍₁₀₎	C ₍₁₁₎	O _(11B)	168(-172)
C ₍₉₎	C ₍₁₀₎	C ₍₁₁₎	O ₍₁₂₎	49(50)
C ₍₉₎	C ₍₈₎	C ₍₁₂₎	O ₍₁₂₎	-55(-51)
O ₍₉₎	C ₍₉₎	C ₍₈₎	C ₍₁₂₎	-70(-71)
C ₍₁₀₎	C ₍₁₁₎	O ₍₁₂₎	C ₍₁₂₎	-62(-59)
O _(11B)	C ₍₁₁₎	O ₍₁₂₎	C ₍₁₂₎	179(161)
C ₍₁₂₎	C ₍₁₃₎	C ₍₁₄₎	C ₍₁₅₎	180(180)
C ₍₁₃₎	C ₍₁₄₎	C ₍₁₅₎	C ₍₁₆₎	-118(-134)
C ₍₁₄₎	C ₍₁₅₎	C ₍₁₆₎	C ₍₁₇₎	56(60)
O ₍₁₅₎	C ₍₁₅₎	C ₍₁₆₎	C ₍₁₇₎	177(-177)
C ₍₁₇₎	C ₍₁₈₎	C ₍₁₉₎	C ₍₂₀₎	174(-55)
O _(1B)	C ₍₁₎	C ₍₂₎	C ₍₃₎	172(179)
C ₍₂₎	C ₍₃₎	C ₍₄₎	C ₍₅₎	177(174)
C ₍₄₎	C ₍₅₎	C ₍₆₎	C ₍₇₎	1(-14)
C ₍₆₎	C ₍₇₎	C ₍₈₎	C ₍₉₎	44(58)
C ₍₇₎	C ₍₈₎	C ₍₉₎	O ₍₉₎	61(56)
C ₍₇₎	C ₍₈₎	C ₍₁₂₎	C ₍₁₃₎	55(60)
C ₍₈₎	C ₍₉₎	C ₍₁₀₎	C ₍₁₁₎	-42(-49)
C ₍₈₎	C ₍₁₂₎	O ₍₁₂₎	C ₍₁₁₎	64(61)
C ₍₉₎	C ₍₁₀₎	C ₍₁₁₎	O _(11A)	(-87)
C ₍₉₎	C ₍₈₎	C ₍₁₂₎	C ₍₁₃₎	-176(-172)
O ₍₉₎	C ₍₉₎	C ₍₁₀₎	C ₍₁₁₎	77(69)
C ₍₁₀₎	C ₍₉₎	C ₍₈₎	C ₍₁₂₎	47(48)
C ₍₁₁₎	O ₍₁₂₎	C ₍₁₂₎	C ₍₁₃₎	-166(-178)
O _(11A)	C ₍₁₁₎	O ₍₁₂₎	C ₍₁₂₎	(68)
O ₍₁₂₎	C ₍₁₂₎	C ₍₁₃₎	C ₍₁₄₎	122(117)
C ₍₁₃₎	C ₍₁₄₎	C ₍₁₅₎	O ₍₁₅₎	120(110)
C ₍₁₅₎	C ₍₁₆₎	C ₍₁₇₎	C ₍₁₈₎	58(71)
C ₍₁₆₎	C ₍₁₇₎	C ₍₁₈₎	C ₍₁₉₎	166(175)

Torsion angles are given to the nearest whole degree. Estimated standard errors in these values range from 2.5 to 7.0° at the C₂₀ chain ends. The β -tail values are given first, followed by the α -tail values in parentheses.

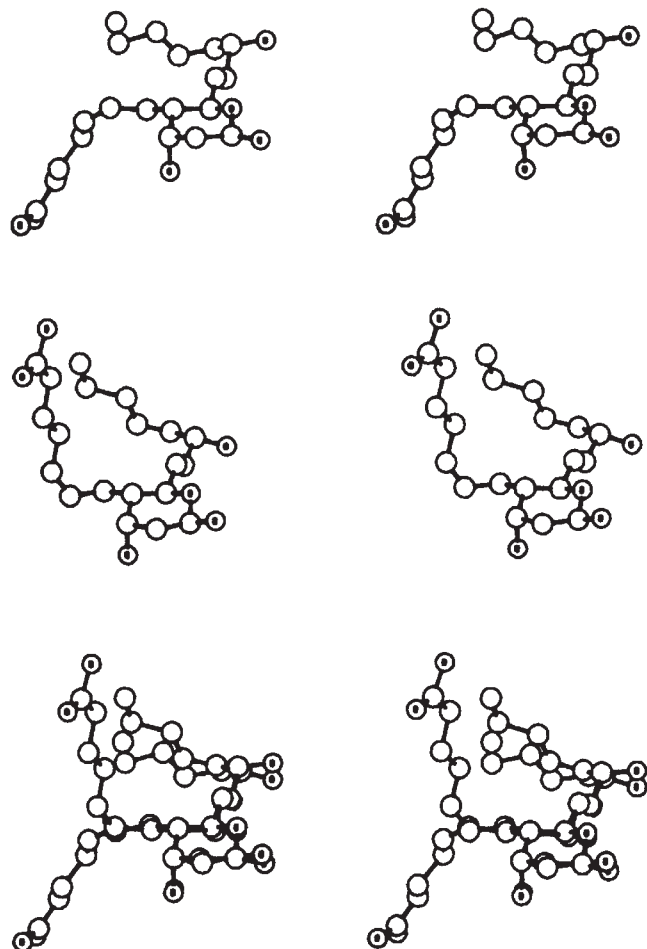


Fig. 2 Stereograms of the α -tail (top) and β -tail (middle) conformers of TXB₂, and a composite (bottom) of the two conformers.

β -tail conformer of TXB₂ also has this conformation. However, the α -tail conformer has a π -anticonformation at C₍₆₎–C₍₇₎. The strong non-bonded interaction between the O₍₉₎ hydroxyl and the C₍₅₎ atom expected for the α -tail conformer has largely been alleviated by closing the C₍₇₎–C₍₈₎–C₍₁₂₎–C₍₁₃₎ torsion angle from the average value of $+86 \pm 5^\circ$ noted for the PGF_{2 α} structures to the observed average of $+58 \pm 4^\circ$ determined for the pyranose side-chain ring junctions of TXB₂.

Whereas the β -tail conformer is identified solely as the 11- β epimer, the α -tail conformer seems to be approximately 65% 11- β epimer and 35% 11- α epimer. Although this corresponds roughly to the 2:1 ratio noted for the unspecified methoxy-TXB₂ epimers resulting from the TXA₂ trapping experiments³, these latter values relate to a different transition state equilibrium from that associated with anomeric equilibrium processes. The two conformers associate as carboxylic acid dimers in the crystal structure.

TXB₂ is a moderately bronchoactive metabolite⁸ which survives degradation by 15-hydroxy-prostaglandin dehydrogenase (PGDH) in the lung^{9,10}. The C₍₁₅₎ hydrogen positions of both conformers may be obstructed from the β -facial approach^{11,12} of the nicotinamide group of NAD⁺ by the π -gauche twist in the ω -chain at C₍₁₆₎–C₍₁₇₎; in addition, the C₍₁₅₎ hydrogen may be obstructed by the α -chain in the β -tail conformer. This may constitute a desirable but not necessarily essential condition for TXB₂ resistance to PGDH metabolism, as the scorpion conformations may be sufficiently different to prevent binding. Although significant amounts of TXB₂ are degraded by PGDH in an antigen-challenged isolated perfusion

preparation of sensitised guinea pig lung tissue¹³, this is not at variance with its resistance in normal physiological conditions. This observation does suggest, however, that thermodynamic preference for the scorpion conformers is not sufficient to prevent less probable but conformationally more compatible forms of the substrate from reacting with the enzyme over prolonged exposure periods.

TXA₂ is physiologically active as a constrictor of coronary arteries¹⁴ and as a potent platelet aggregator³ in thrombus formation, whereas TXA₁, derived from PGH₁, lacks both activities. This has implied the chemical importance of *cis*- Δ^5 unsaturation for hormonal recognition in platelets and vascular smooth muscle^{15,16}. The present study calls attention to the conformational importance of the *cis*- Δ^5 bond to produce α -tail and β -tail scorpion conformers. It remains to be shown whether the structure of TX receptors fits either of these conformational models and whether the conformational requirements for platelet and smooth muscle receptors are different.

The TXB₂ samples used in this study were synthesised^{17–19} at the Upjohn Company, and provided by Drs Bundy and Pike. This work was supported by grant no. HL-15378 awarded by the National Heart, Lung and Blood Institute, DHEW.

Received 2 January; accepted 6 August 1979.

- Andersen, N. H., Ramwell, P. W., Leovey, E. M. K. & Johnson, M. in *Advances in Prostaglandin and Thromboxane Research* Vol. 1 (eds Samuelsson, B. & Paoletti, R.) 271–289 (Raven, New York, 1976).
- Hamberg, M. & Samuelsson, B. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3400–3404 (1974).
- Hamberg, M., Svensson, J. & Samuelsson, B. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2994–2998 (1975).
- Hanessian, S. & Lavalley, P. *Can. J. Chem.* **55**, 562–565 (1977).
- Nyburg, S. *Acta crystallogr. B* **30**, 251–253 (1974).
- Chidester, C. G. & Duchamp, D. J. *American Crystallographic Association Spring Meeting*, Berkeley, California, Abstr. A6 (1974).
- Langs, D. A., Erman, M. G. & DeTitta, G. T. *Science* **197**, 1003–1005 (1977).
- Wasserman, M. A. & Griffin, R. L. *Eur. J. Pharmac.* **46**, 303–313 (1977).
- Kindahl, H. *Prostaglandins* **13**, 619–629 (1977).
- Roberts, L. J., Sweetman, B. J., Morgan, J. L., Payne, N. A. & Oates, J. A. *Prostaglandins* **13**, 631–648 (1977).
- DeTitta, G. T. *Science* **191**, 1271–1272 (1976).
- DeTitta, G. T., Langs, D. A. & Edmonds, J. W. *Biochemistry* **18**, 3387–3391 (1979).
- Dawson, W., Boot, J. R., Cockerill, A. F., Mallen, D. N. B. & Osborne, D. J. *Nature* **262**, 699–702 (1976).
- Piper, P. J. & Vane, J. R. *Nature* **223**, 29–35 (1969).
- Needleman, P., Minkes, M. & Raz, A. *Science* **193**, 163–165 (1976).
- Falardeau, P., Hamberg, M. & Samuelsson, B. *Biochem. biophys. Acta* **441**, 193–200 (1976).
- Nelson, N. A. & Jackson, R. W. *Tetrahedron Lett.* **37**, 3275–3278 (1976).
- Kelly, R. C., Schletter, L. & Stein, S. J. *Tetrahedron Lett.* **37**, 3279–3282 (1976).
- Schneider, W. P. & Morge, R. A. *Tetrahedron Lett.* **37**, 3283–3286 (1976).

Corrigendum

In the article 'Radio studies of the double QSO, 0957 + 561A,B' by G. G. Pooley *et al.* *Nature* **280**, 461–464, in Table 2 the entry for radio component A should indicate a size of 1 arc s. In the equation below Fig. 3, for ... independent of data ..., read ... independent of β .

Erratum

In the article 'Codon–anticodon interaction at the ribosomal peptidyl site' by R. Lührmann *et al.* *Nature* **280**, 423–424, Figures 1 and 2 are transposed. Figure 1 shows dependence on [Mg²⁺].

reviews

Tour of statistical history

Stephen M. Stigler

The History of Statistics in the Seventeenth and Eighteenth Centuries, against the changing background of intellectual, scientific and religious thought. Lectures by Karl Pearson given at University College, London, during the academic sessions 1921-1933. Edited by E. S. Pearson. Pp. 744. (Charles Griffin: High Wycombe, UK, 1979.) Paperback £7.50.

KARL PEARSON combined apparently boundless energy with a wide-ranging enquiring intelligence and indomitable perseverance to produce a remarkable career as a philosopher, physicist, mathematician, statistician, historian, editor, biologist and teacher. Pearson will doubtless be best remembered as a statistician. His first statistical publication was a letter to *Nature* on asymmetrical frequency curves (published October 26, 1893). Over the succeeding decade he produced such widely known techniques as the χ^2 test, and he was primarily responsible for the early development of correlational analysis. The fact that such a basic term as "standard deviation" was first introduced by Pearson is testimony to the ubiquity of his influence upon subsequent statistical thought. When Pearson died in 1936, at the age of 79, he left behind a published legacy so extensive that a medium-sized hardcover book was required for his bibliography alone. He also left behind a large archive of correspondence and unpublished lecture notes, and it is from this latter source that this present volume has been drawn.

Before Pearson's emergence as a researcher in statistical methods in 1893, he had demonstrated an interest and ability in the history of science through his editing and completion of the second volume of Isaac Todhunter's *A History of the Theory of Elasticity and of the Strength of Materials from Galilei to the Present Time* (1893). After the War his thoughts returned to historical matters, as he sought to complete a massive biography of Francis Galton, and from 1921 to 1933 he gave a course of lectures at University College, London, on the history of statistics. Fortunately for all of us the notes of these lectures have survived and found the perfect editor — Pearson's son Egon S. Pearson, himself justly celebrated for both his innovative research in statistics and his insightful accounts of its history.

Karl Pearson was known as a spell-



Pearson with the Brunsviga adding machine, 1910

Courtesy University College, London, UK

binding lecturer, and a good sense of his pace, method and vocal energy comes across on these pages. Apparently Pearson's habit was to draft his lectures *in extenso*, as if for reading, and it is these drafts which have survived. Thus, the present work is essentially a verbatim recording of the lectures of a master lecturer, a connected and detailed account of its topic at a pace (and with the necessary digressions) to hold a listener's attention. For the most part this is not the dry chronological account that so many histories of this field become, but it is Karl Pearson speaking to us. Sometimes he is earnest and scolding, as when he remonstrates with d'Alembert for a failure to test his philosophical views experimentally (page 542); sometimes he goes off on interesting, if irrelevant digressions, as when he disputes statements that Lagrange's first marriage had been incompatible (page 578). Occasionally there are flashes of humour, as when, after explaining how Lagrange loved to abandon himself to mathematical thought while listening to music, Pearson wrote: "Luckily 'Jazz' was not then invented, or the result might have made him a circle-squarer, rather than one of the greatest of mathematicians" (page 576).

These lectures cover the period from John Graunt's 1662 *Observations on the London Bills of Mortality* to Laplace's 1812 *Théorie Analytique des Probabilités*, with considerable emphasis on biographical detail and personality. The coverage reflects Pearson's interests and

his taste for philosophy. There is extensive discussion of early work by the political arithmeticians (including estimation of population, construction of life tables, and inference about the sex ratio), but very little on the handling of astronomical or geodetical data. Pearson took a particular fancy to individuals with an interest in both mathematics and theology, and Richard Price receives the benefit of 52 pages of discussion. In his discussion of the later mathematicians, Lagrange and Laplace, Pearson's attention is more focussed on mathematical detail, in the manner of Todhunter's *A History of the Mathematical Theory of Probability*. But even here he exults when he encounters an antecedent to one of his own results. His discovery (page 599 *et seq.*) that Lagrange had almost derived the χ^2 distribution in about 1770 shows us both the excellence of Lagrange's mathematical skills and Pearson's own tendency to elevate mathematical form above statistical concept.

As a history of statistics, the book is not without its defects, and many are pointed out by the editor in his preface. There are errors of fact (though few misprints) that Pearson would likely have caught had he revised the lectures himself, and at times the length of the exposition makes for heavy going. The views presented of past accomplishments are dated by being rooted in Pearson's own work, and the background of religious thought and political movement that

Pearson gives in many extensive digressions is not, I think, successfully integrated into the discussion of the development of statistical methods. Bibliographical details are generally given accurately, but references to even those secondary sources available to Pearson are scarce. But these are all minor problems compared with the rare opportunity we have been given to sit at the feet of this energetic and learned professor as he takes us on a personalised tour of statistical history, and we often learn as much in the process about Pearson as we do of his subject. Here Pearson speaks of a table of Simpson's as "probably faked" (page 387); there Pearson laments that the brains of Lagrange and Laplace were not preserved for later study (page 585); and there Pearson faults Euler for failure to compile numerical tables "which would energise

the symbols" (page 245). Whether we are annoyed by his opinionated view of history or enthralled by his enthusiasm, we cannot fail to be fascinated and grateful for this chance to witness these lectures.

The editor, E. S. Pearson, has done an excellent job of piecing together the different versions of several lectures, furnishing a useful index and many notes, and cutting wisely where cutting was needed. The publishers must be congratulated for keeping the price at what must, for a volume of this size, be considered low. We may hope that this encourages many scientists (not only statisticians) to make the acquaintance of the extraordinary Karl Pearson. □

Stephen M. Stigler is a Professor in the Department of Statistics at the University of Chicago, Illinois.

Nutrition planning

Nutrition and National Policy. Edited by Beverley Winikoff. Pp. 580. (MIT Press: Cambridge, Massachusetts, and London, UK, 1979.) £15.75.

THIS book is a monument to nutrition planning, the latest enthusiasm to sweep the nutrition profession as it looks desperately for some effective action with which to improve the health problems of developing countries.

Nutrition planning is a controversial approach the more coherent advocates of which argue that protein energy malnutrition in children must not be seen simply as a disease entity demanding clinical treatment. It is a disease of the poorest of the poor, a symptom of deprivation, the product of a lack of wealth and welfare which can only be eliminated by action at the economic level. As the object of development planning is to help reduce these same inequalities, three important consequences follow. Nutritional status will be improved by successful development strategies, and the improvement in nutritional status must be one of the objectives of a good development policy. Conversely, the extent of malnutrition serves as an

indicator of the effectiveness of those programmes.

This approach is one that I and many of my colleagues find acceptable. Many others do not agree, finding it unrealistic or unnecessary. These are legitimate differences in opinion, a source of mutual respect and a fruitful academic dialogue.

Unfortunately, although nutrition planning is an exciting theory, it is also a popular idea. The concept of nutrition planning has not been accepted but the image has, and this has become the portmanteau justification of tarted up versions of the same tired old theories that failed in the past. Sadly, ideas change slowly: they are simply renamed and thereby extended.

From all this comes, I assume, this book — almost 600 pages of mostly unnecessary prose with all too few passages that are worth reading. Two-thirds of the book are devoted to 'case studies' of national nutrition policies largely by the representatives of various developing countries. Apart from a brilliant study of malnutrition in Chile by G. Solimano and P. Hakim, which is a précis of their book (*Development, Reform, and Malnutrition in Chile*, MIT Press; for review, see *Nature*, 276, 542; 1978) on this topic, these national presentations are mediocre indeed. If nutrition students presented, as their view on nutrition planning, the hotch potch of miscellaneous data that seem to satisfy some of these planners they would fail their degrees. Why then are they included here? One possible reason lies in the origins of the book, which is the edited proceedings of a conference funded by the Rockefeller Foundation which took place at a Rockefeller mansion in Bellagio, Italy in 1975.

There is a growing tendency in international jamborees devoted to such areas of nutrition to invite participants

not by eminence, or relevance, but by race. This may be politically expedient, but it is bad for science, and ultimately insulting to those who only get invited as statutory representatives of the third world. I do not know how far this consideration governed the structuring of this meeting, but, distasteful though the idea is, it would explain the quality of the papers selected, which add little to human understanding and detract from the science of nutrition.

It seems to me that nutritionists in government should at least have details of the prevalence of malnutrition in their countries, and changes in that rate with development policy. No such luck. If these authors are really reporting the understanding of nutrition planners in their countries, then either systematic data is not being collected, or no-one exploits it in the way that is surely a precursor to its use in the political decision making process.

We do not need them to report over and over again that malnutrition occurs in children, or even that the poor are nutritionally worse off than the rich. Such information is not in doubt. It is clear that the poor suffer. What is also clear is that even at the worst not all the poor are malnourished. Nutrition planning will only proceed effectively when the political decision makers are given some idea of which socially or ecologically definable groups are at greatest risk, so that social policy can then be designed to help those groups.

In this book economic and nutritional information is presented in abundance, but it is presented neither within a framework that presupposes such action nor in the form that can be used by the reader to develop his own theories. I cannot imagine a politician being interested in such ramblings. These are not case studies they are travelogues compounded of largely uncollated nutritional and economic data, and peppered with politically tactful remarks.

When the book gets on to generalisations it improves somewhat, but still says little at great length. Towards the end it is suggested that Hong Kong, Singapore, South Korea and Japan have rapidly improved nutritional status as a result of income redistribution. I am not convinced, but given that it is even tenable why doesn't the Rockefeller Foundation now fund a less lavish session to discuss the situation in these countries. If it were done coherently, those of us interested in this proposition could see whether given the political will, economics alone was the answer to nutritional problems. That at least would provide a test of the nutrition planning hypothesis in a way that the present volume notably fails to do.

John Rivers

John Rivers is Lecturer in the Department of Human Nutrition at the London School of Hygiene and Tropical Medicine, London, UK.

● In the review of *Atomic Absorption Spectroscopy* (see *Nature*, 280, 705; 1979), the author's name was incorrectly quoted. This should have read M. Slavin.

● *Origins: What New Discoveries Reveal about the Emergence of Our Species and Its Possible Future* (for review, see *Nature*, 270, 108; 1977), by Richard Leakey and Roger Lewin, has been published in paperback by Macdonald and Jane's (London, UK).

Atlas of economic mineral deposits

Atlas of Economic Mineral Deposits. By C.J. Dixon. Pp. 143. (Chapman and Hall: London, 1979.) £35.

MINERAL deposits of economic grade represent in all cases abnormal concentrations of the contained elements resulting from unusual combinations of geological processes. This is equally true whether the deposits are of metalliferous ore or industrial minerals, and such deposits are dispersed very thinly through the outer crusts of the Earth. Their bulk is usually small compared with that of the rocks with which they are associated. These fundamental points are implied in Colin Dixon's original and interesting *Atlas*, but they might have received more specific emphasis. His book deals with solid, non-combustible mineral deposits, and bears upon energy sources only in the case of uranium minerals. The deposits with which he is concerned supply, however, most of the raw materials for human artefacts, the manufacture of which is the purpose of the greater part of industry. The *Atlas* is intended as a reference work and a source of descriptive material to accompany a university course. The student of geology whose course includes (regrettably!) no discussion of the useful minerals would also find this a good text to read.

In nine pages of introduction set in three-column format Dixon deals with the generalities of the subject, including prospecting, exploitation, beneficiation and, briefly, the economics of the business; also the scientific aspects including petrology, morphology, and distribution, concluding with a neat summary of the history of genetic interpretation. Some geologists might have been tempted to generalise more about the structural control of ore shoots — the area where the application of geological reasoning has contributed most to industry — but this emerges well enough from the detailed descriptive pages that follow. Here, instead of attempting a comprehensive survey, forty representative major mineral-producing districts are illustrated and summarised, both aspects being ably presented. A very brief list of references accompanies each description forming a tailpiece to the accounts, each of which occupies a double folio of the 240 x 400 mm format. In most cases the drawings, on which a uniform system of symbols is used, are simplified from the published originals, often with great advantage in bringing out the essential points under discussion. Almost all teachers of the subject would agree with the author's contention that the subject needs visual presentation.

The major grouping is fourfold: (1) deposits in environments at the Earth's surface; (2) in sedimentary rocks; (3)

associated with felsic magmatic environments; (4) in basic and ultrabasic magmatic rocks. This implies only a loose genetic classification, and wisely so, as it avoids the still-controversial question as to whether the mineral concentrations in sediments were formed in the normal course of sedimentation or not. As the author remarks, a group of French and other continental workers feel strongly that in fact they are so formed, but I am glad that he was not able to think of an English translation of *la gîtologie* for this term seems to me to be very closely associated with this school, the conclusions of which many of us find quite unacceptable in the light, for example, of inclusion homogenisation and stable isotope studies. The compromise proposal which says that minerals concentrated during some later diagenetic episode are sedimentary in origin also seems unsatisfactory to some of us. Mr Dixon wisely avoids confrontation by saying that the terms syngenetic and epigenetic are difficult to apply. His grouping brings together the exogenous deposits in (1) and it is open to those who wish to do so to think that the deposits in groups (2) to (4) derived their energy of formation from processes internal to the Earth's crust and mantle.

The forty examples selected give a good

picture of the range of different geological environments in which valuable deposits may become concentrated. The processes involved probably varied as widely as the environments in which they operated. If there is any common factor, surely it is the efficacy of brine as a solvent of metals and other useful elements, whether as sea-water passing through hot basic lavas extruded on the ocean floor, or connate fluid sucked from sediments through cooling felsic magmas, or hypersalines forced out of deep consolidating sedimentary basins. But no single mechanism will cover all possibilities and we should agree with the author that the multiple hypothesis remains unavoidable.

A final section contains five world-distribution maps, and appendices on mineral names, units of measurement and stratigraphic nomenclature. But for the price, I would recommend every student of geology to acquire this book. The hope should be expressed that the unusual shape will not lead to it becoming hidden in obscure corners of college libraries.

Kingsley Dunham

Sir Kingsley Dunham was Director of the Institute of Geological Sciences from 1967-75.

Digging into Wittgenstein

Ludwig Wittgenstein: Remarks on the Foundations of Mathematics. Revised edition. Edited by G.H. von Wright, R. Rhees and G.E.M. Anscombe. Translated by G.E.M. Anscombe. (MIT Press: Cambridge, Massachusetts, and London, 1978.) \$27.50.

THIS is a considerably expanded version, in English only, of the book first published under the same title, in English and German, in 1956. It consists, as the editors explain in their Preface, of selections from Wittgenstein's manuscript writings on the philosophy of mathematics from 1937 to 1944, together with nine pages written in 1933-34. It is divided into seven parts, with three appendices to Part I, each part being taken from a different manuscript; but, as the editors make clear, only Parts I and VI reproduce Wittgenstein's text *in toto*, the others constituting selections made by the editors from more extensive writings. The editors express "the opinion that the time has not yet come to print the whole of Wittgenstein's MSS on these and other topics", but give no reasons.

It is now approaching thirty years since Wittgenstein's death; and it is only quite recently that what might be called the philosophical community has come to real grips with the most difficult and profound of his ideas. In that time, much chaff has

blown away. The heady excitement of logical positivism, and, after it, of Oxford ordinary language philosophy, can be looked back on only with some embarrassment as intoxication with crudely oversimplified ideas. Wittgenstein's work remains; undeniably, now, that of one of those few philosophers who will be read by all future generations. It is by far the richest twentieth-century source of philosophical ideas, which it will take us more decades yet properly to apprehend and to absorb; despite the difficulty with which his work presents the reader, there is nothing that is likely to be more rewarding. The philosophy of mathematics was one of his earliest and most persistent pre-occupations. Originally he planned to include a lengthy discussion of it in the *Philosophical Investigations*, but, being dissatisfied with what he had written, cut almost all of it out; this material comprises Part I of the present book. He was evidently never satisfied with the work he had done in this area of philosophy, and, as the editors remark, did not return to it after April, 1944. The present edition is thus in no sense a finished treatment, or even a unitary book, any more than was the original one: it is a selection from seven distinct pieces of writing by Wittgenstein, with none of which he was content. For all that, it demands the most thorough attention from anyone interested in his philosophy, because the subject occupied so important a place in his thought. He obviously believed that his basic philosophical ideas had profound consequences

for the philosophy of mathematics. Precisely what he took these consequences to be is something that the reader has to work very hard to discover.

Wittgenstein's mode of composition was unique. He filled notebooks with 'remarks', a few lines in length, which he continually revised; he would also constantly rearrange the 'remarks' in different sequences to form a disjointedly continuous text. Sometimes the import of the individual remark is luminously clear, sometimes quite obscure: but the method of composition obviously risks leaving the sequence of thought difficult to follow. That it need not have this effect is apparent from the *Investigations*, the only product of his later period that Wittgenstein regarded as a (nearly) finished book: there the final product of years of revision and rearrangement provides a largely continuous treatment of a range of topics, and, given close concentration on the reader's part, expresses with considerable clarity Wittgenstein's conclusions concerning them. Unfortunately, such an outcome to so strange a mode of composition was usually attained only at the final stage: the present book contains seven different sequences of remarks, five of them with omissions made by the editors, and none of them representing a final stage. It is difficult to convey the obstacles which this presents to comprehension without extensive quotation, but perhaps the following brief sequence from Part IV may be regarded as not untypical:

37. A human being is imprisoned in a room, if the door is unlocked but opens inwards; he, however, never gets the idea of pulling instead of pushing against it.

38. When white turns black some people say "Essentially it is still the same"; and others, when the colour turns a shade darker: "It is completely different".

39. The proposition ' $a = a$ ', ' $p \supset p$ ' ['If p , then p '], 'The word "Bismark" has 8 letters', 'There is no such thing as reddish-green', are all obvious and are propositions about essence: what have they in common? They are evidently each of a different kind and differently used. The last but one is the most like an empirical proposition. And it can understandably be called a synthetic *a priori* proposition.

It can be said: unless you put the series of numbers and the series of letters side by side, you cannot know how many letters the word has.

Of course; no-one can be expected to understand the drift of such a sequence torn from its context: but it illustrates reasonably well the difficulties with which the reader has to contend. We have first two striking illustrations: but what exactly they are intended as illustrations of has to be gathered from the whole context in which they are embedded, bearing in mind that, if Wittgenstein had revised his text further, they might have been placed in a quite different immediate context, and bearing in mind also that what is printed here is only a selection, and the editors have not indicated where the omissions occur. There follows, in § 39, what can only be described as a series of musings. None of

the observations it contains can possibly be gainsaid: our difficulty is to know what *force* they have. The four propositions are indeed obvious: what connection is this supposed to have with their bearing 'about essence'? What relation do such propositions have to others that are equally 'about essence' but are not at all obvious? In saying that they are 'about essence', is Wittgenstein merely borrowing a term from another philosophy in order to express the *a priori* character which anyone would be likely to see as attaching to all four propositions, or does he have a sharp idea of essence? Do the propositions have *more* in common than what has already been stated, namely that they are obvious and are about essence? They are evidently all different, and differently used: but how great are we supposed to see these differ-

ences as being? The third proposition may understandably be called synthetic *a priori*: but can it justifiably be so called? To obtain any philosophical illumination from a passage like this, the reader must answer all these questions and more: and he must answer them not only to his own satisfaction, but in the manner that it is plausible that Wittgenstein intended them to be answered. As Frege said of his own posthumously published writings, not all of what is contained in this book is gold, but there is gold in it: but, in Wittgenstein's case, the mining operation is particularly arduous.

Michael Dummett

Michael Dummett is Fellow of All Souls College, Oxford, UK.

Geochemical handbook

Handbook of Geochemistry. Vol. 2, Part 5: Elements La (57) to U (92). Edited by K.M. Wedepohl. Pp. 1546. (Springer: Berlin, Heidelberg and New York, 1979). Loose-leaf binder DM690; \$379.50.

FOUR years have elapsed since the publication of Part 4 of the *Handbook of Geochemistry*, the longest gap in the production schedule, which has extended almost over a decade from 1969 to 1978. Part 4 is also by far the largest to appear, comprising more than 1500 pages. The original intention of the editorial board was that Volume 2 of the *Handbook* would ultimately contain approximately 2000 pages in four instalments, with publication completed several years ago. The much larger compilation and considerably extended publication time no doubt reflect many factors, not least the rapid growth of geochemical literature and the fact that all authors contributing to the project did not work at the same rate.

The *Handbook* was probably never intended, even in early days, for the individual purchaser. It is expensive to buy, now costing on the subscription price \$785.60 (DM 1428.40), including Volume 1, which contained a series of chapters on geochemistry, geophysics and cosmochemistry. Volume 2 is, however, the principal expense. The dollar price for each part has, not surprisingly, risen over the years, but the price in German marks for each part has been held to the level at time of publication.

What does the purchaser receive for his money? In Part 5 he is presented with the pages to complete a large number of the chapters on individual elements, some of which were begun almost a decade ago. Several of these elements, notably phosphorus, sulphur, titanium, cerium, manganese, cobalt, strontium and molyb-

denum, are important geochemically, and virtually no information has been available for these elements until now in the *Handbook*. Although this lack of data has at times been frustrating, at least there is now the compensation that these chapters are the most up-to-date in the book. Many new pages carrying amendments that include additional information and the correction of scientific and printing errors, are also supplied with Part V.

Of the *Handbook* as a whole, once rearrangement of the entire loose-leaf system — not really a drawback, despite editorial concern — has been completed, the text represents the most comprehensive compilation of geochemical data so far attempted in one publication. The *Handbook* is logical in lay-out, with a high publishing standard of text (and the many tables and diagrams) maintained throughout. As more than 100 authors have been involved in the project, it is inevitable that the style, presentation and depth of coverage between, and even within, chapters will vary. The reader will have to accept this lack of uniformity, just as he will have to accept that much of the text is not up-to-date — the first manuscript was received in 1965. The editorial board now accepts that revision of earlier text to include more recent information is impossible, whatever its members might have felt individually or collectively in the euphoria at the start of the project. They must be relieved it is all over and it is perhaps unlikely that any one of them will embark again on such a large-scale project. The editors and authors should receive the congratulations and thanks of geochemists, first for successfully completing the *Handbook* and second for providing such a valuable authoritative base of data that will make the geochemist's professional life so much easier.

D.G. Murchison

D.G. Murchison is Professor of Organic Petrology and Past Reader in Geochemistry at the University of Newcastle-upon-Tyne, UK.